



USAID
DU PEUPLE AMERICAIN

SENEGAL

ETUDE SUR LA MAITRISE DU VOCABULAIRE ORAL EN FRANÇAIS DES ELEVES DE CI, CP ET CEI AU SENEGAL

USAID/LECTURE POUR TOUS

Soumis:	3 Fevrier, 2020
Resoumis:	21 Avril, 2020
Numéro du Contrat :	AID-OAA-I-14-00055/AID-685-TO-16-00003
Date de Démarrage et de Fin des Activités :	26 octobre 2016 -10 juillet 2021
Montant total de la Subvention :	\$71,097,573.00
Représentante de l'Agent Contractant :	Kadiatou Cisse Abbassi

Soumis par: **Chemonics International**
Sacré Coeur Pyrotechnie Lot No. 73, Cite Keur Gorgui
Tel: 221 78585 66 51
Email: LecturePourTousPMU@chemonics.com

ETUDE SUR LA MAITRISE DU VOCABULAIRE ORAL EN FRANÇAIS DES ELEVES DE CI, CP ET CEI AU SENEGAL

Contrat sous AID-OAA-I-14-00055/AID-685-TO-16-00003

USAID/Lecture Pour Tous

AVERTISSEMENT

Les avis exprimés par l'auteur de cette publication ne reflètent pas nécessairement ceux de l'Agence des États-Unis pour le Développement International ou du Gouvernement des États-Unis.

SOMMAIRE

RESUME EXECUTIF	6
INTRODUCTION	10
1. APERCU DU CADRE THEORIQUE.....	11
2. METHODOLOGIE	16
2.1. Présentation de l'enquête	16
2.2. Présentation de l'échantillon	17
2.2.1. Les critères de sélection	18
2.2.2. Les caractéristiques des élèves de l'échantillon	18
2.3. Protocole d'évaluation	26
2.3.1. Les tests développés.....	26
2.3.2. Les caractéristiques des mots intégrés dans les tests	27
3. RESULTATS.....	29
3.1. Les résultats obtenus aux items des épreuves de vocabulaire et de lecture	29
3.1.1. Épreuves de vocabulaire oral en français	29
3.1.2. Épreuves de prélecture et de lecture.....	33
3.2. Les scores obtenus aux épreuves de vocabulaire et de lecture.....	34
3.2.1. Vocabulaire oral en français	35
3.2.2. Prélecture et lecture	39
3.3. Comment expliquer la variabilité des scores des élèves ?	43
3.3.1. La démarche d'analyse.....	43
3.3.2. Résultats au test de vocabulaire oral en français.....	45
3.3.4. Résultats aux tests de prélecture et de lecture	65
4. SYNTHESE ET DISCUSSION	75
4.1. Rappel des hypothèses	75
4.2. Evaluation des hypothèses sur le niveau du langage oral en français et de lecture....	76
4.2.1. Épreuve de vocabulaire (du mot au groupe de mots).....	76
4.2.2. Épreuve de lecture.....	77
4.3. Autres facteurs ayant une incidence sur les capacités évaluées	78
4.3.1. Résultats contradictoires à propos de l'environnement géographique	78
4.3.2. A propos des activités scolaires	79
4.3.3. A propos du niveau de diplôme des enseignants	80
4.3.4. A propos des modalités d'organisation de la classe.....	81
4.4. Implications des résultats de l'étude pour les politiques éducatives	82
4.4.1. Implications des résultats concernant les questions au centre de l'étude	82

4.4.2. Autres implications pour les politiques éducatives.....	83
4.5. Quelques limites de la présente étude.....	85
4.6. Pour conclure.....	85
5. BIBLIOGRAPHIE.....	87
6. ANNEXES.....	91
6.1. Protocole d'évaluation.....	91
6.2. Questionnaire Élève.....	99
6.3. Questionnaire Parent d'élève.....	100
6.3. Questionnaire Enseignant.....	102

AVANT-PROPOS ET REMERCIEMENTS

Pour améliorer la qualité des apprentissages des élèves, le Gouvernement du Sénégal s'est engagé dans une réforme de l'enseignement de la lecture dans les premières années du cycle primaire. Cet effort bénéficie d'un appui de l'USAID à travers le programme Lecture Pour Tous (LPT ; 2016-2021). De façon plus spécifique, ce programme soutient l'introduction de nouvelles méthodes d'enseignement-apprentissage de la lecture et notamment l'utilisation pour cet enseignement des trois langues nationales que sont le pulaar, le seereer et le wolof pour les élèves des trois premières classes (CI, CP et CEI) des écoles élémentaires publiques et des daaras de sept régions du pays (Diourbel, Fatick, Kaffrine, Kaolack, Louga, Matam, Saint-Louis¹).

Le programme Lecture Pour Tous comprend un volet recherche, qui vise à mener des recherches opérationnelles pour informer les activités du programme et notamment l'appui aux politiques. Cette étude s'inscrit dans ce dispositif. Elle vise d'une part, à évaluer le niveau de maîtrise du vocabulaire oral en français des élèves en début de primaire et dans les premières années du cycle, et d'autre part, à appuyer le Ministère de l'Education Nationale (MEN) pour définir le moment et les modalités d'introduction du français dans l'enseignement de la lecture. Dans cette perspective, les informations utiles pour mener à bien ce travail ont nécessité la construction de différents instruments et leur administration à différents acteurs.

La construction des instruments de l'enquête a débuté en septembre 2018. Le cadre de référence de l'étude et le protocole d'évaluation des compétences des élèves ont été définis par Liliane Sprenger-Charolles (Centre National de la Recherche Scientifique (CNRS) / Université d'Aix-Marseille), tandis que les questionnaires contextuels ont été construits par Adeline Seurat (Institut de Recherche sur l'Education (IREDU) / Université de Bourgogne). Tous les instruments ont été ensuite discutés et validés par une équipe de l'Institut National d'Étude et d'Action pour le Développement de l'Education (INEADE). L'échantillon de l'enquête a été mis au point par l'équipe Monitoring, Evaluation and Learning (MEL) du programme Lecture Pour Tous.

La formation des enquêteurs s'est déroulée en novembre 2018 à Kaolack. L'administration de l'enquête s'est déroulée au cours du même mois dans trois Inspections de l'Éducation et de la Formation (IEF) : l'IEF de Gossas dans l'Inspection d'Académie (IA) de Fatick, l'IEF de Kaolack Commune dans l'IA de Kaolack et l'IEF de Kaffrine dans l'IA de Kaffrine. Elle a été menée sous la supervision d'Ibrahima Cissé et de Mamadou Dramé pour le programme Lecture Pour Tous, et d'Alioune Cissé, d'El Hadji Sonko et de Birama Touré pour l'INEADE. Assana Diop (Lecture Pour Tous) a préparé les tablettes pour la collecte des données et a constitué les fichiers en décembre 2018.

Enfin, les travaux d'analyse des données et de rédaction du rapport ont été réalisés par Adeline Seurat et Liliane Sprenger-Charolles entre décembre 2018 et mai 2019. Une équipe technique de l'INEADE, de la DEE et de la DALN a participé à un atelier de 5 jours en avril 2019 pour regarder les premières analyses des données et pour formuler des résultats préliminaires de l'étude.

¹. La région de Saint-Louis bénéficie seulement de l'appui technique du programme Lecture Pour Tous, tandis que les autres bénéficient d'un appui technique et financier.

RESUME EXECUTIF

Une question cruciale pour le programme Lecture Pour Tous, qui a pour objectif de faciliter l'apprentissage de la lecture au Sénégal grâce à l'utilisation de nouvelles méthodes d'enseignement dans le contexte des langues nationales (LN), est de déterminer le moment le plus propice pour introduire l'enseignement de la langue seconde (L2) sur la base de données scientifiques pertinentes. Le programme Lecture Pour Tous a donc été chargé, selon son mandat contractuel avec USAID, d'une évaluation du niveau de vocabulaire oral en français des élèves dans les trois premières classes du primaire. Les informations présentées dans ce document ont pour but d'informer les choix techniques et politiques du MEN dans le cadre de leur modèle d'enseignement bilingue, y compris ceux concernant l'enseignement de la L2 à l'oral, afin de faciliter l'introduction de la lecture dans cette langue et de permettre le transfert des acquis de cet apprentissage en LN.

Pour cette étude des tests visant à évaluer le niveau de vocabulaire oral des élèves de CI, CP et CEI, ainsi que leur niveau de prélecture (CI-CP) et de lecture (CEI), ont donc été élaborés. Des questionnaires visant à recueillir des informations sur les facteurs linguistiques, sociologiques et pédagogiques susceptibles d'influencer les résultats des évaluations ont également été construits et administrés aux élèves, à leurs parents, ainsi qu'aux enseignants que les élèves ont eus l'année qui a précédé l'étude (ceux de CI pour le CP, ceux de CP pour le CEI). L'équipe de Lecture Pour Tous a démarré la construction des instruments de l'enquête au mois de septembre 2018. Un processus d'échange et de revue conjointe a été menée avec l'INEADE et la participation de la DALN et de la DEE, afin de stabiliser les instruments d'enquête. Ce processus a abouti à la définition de quatre instruments : un protocole d'évaluation du vocabulaire français oral des enfants sénégalais et de leurs capacités de lecture, assorti de trois questionnaires. En parallèle, Lecture Pour Tous a coordonné avec le personnel de l'INEADE la construction de l'échantillon pour cette enquête. Pour des raisons logistiques, trois régions proches les unes des autres ont été choisies : celles de Fatick, Kaolack et Kaffrine. Parmi elles, trois Inspections d'Education et de Formation (IEF) ont ensuite été sélectionnées de façon aléatoire : l'IEF de Gossas dans l'Inspection d'Académie (IA) de Fatick, l'IEF de Kaolack Commune dans l'IA de Kaolack et l'IEF de Kaffrine dans l'IA de Kaffrine. La construction de l'échantillon de l'enquête a ensuite cherché à respecter (i) la répartition du nombre d'élèves scolarisés entre ces trois IEF, puis au sein de chaque IEF, la répartition (ii) selon le milieu de résidence (urbain / rural) et (iii) selon la langue choisie pour l'école et enfin, (iv) la répartition des écoles selon l'année de mise en place du programme (an 1 ou an 2). Une fois ces critères appliqués, un certain nombre d'écoles ont été sélectionnées selon des modalités aléatoires. Lecture Pour Tous a ensuite collaboré étroitement avec l'INEADE pour réaliser un pré-test des outils à Dakar et procéder à la formation de 14 enquêteurs et 3 superviseurs, du 05 au 9 novembre 2018. La collecte de données a eu lieu directement après la formation, durant 10 jours. Au total, l'échantillon compte 1 183 élèves de CI (388), de CP (399) et de CEI (396) répartis dans les trois IEF. Ces différentes données ont ensuite été consolidées et fusionnées dans une même base afin de disposer des résultats de chaque élève à chaque item du protocole d'évaluation, de connaître ses caractéristiques et celles de ses parents, ainsi que, pour ceux de CP et de CEI, de disposer d'informations sur ce qu'ils ont fait en classe l'année précédente. Les premières analyses ont fait l'objet d'une revue conjointe avec l'INEADE afin d'alimenter le présent rapport.

Les travaux de recherche dans le monde entier indiquent que, dans une langue donnée pour bien comprendre un message écrit, il faut être capable, d'une part, de comprendre le même

message à l'oral et, d'autre part, avoir maîtrisé le décodage (évalué par la lecture de mots inventés) ainsi que la procédure lexicale d'identification des mots écrits (évaluée par la lecture de mots fréquents). La rapidité de la maîtrise du décodage dépend de la régularité de correspondances graphème-phonème (CGPh) dans la langue dans laquelle s'effectue l'apprentissage et celle de la procédure lexicale du niveau de vocabulaire oral de l'apprenant. Les travaux de recherche indiquent également que la maîtrise du décodage est un puissant mécanisme d'auto-apprentissage, ce qui s'explique par le fait que, quand les CGPh les plus régulières ont été acquises, il devient plus facile d'apprendre celles qui le sont moins et, en fonction du niveau de vocabulaire, de mettre alors en place la procédure lexicale d'identification des mots écrits.

Ces explications permettent de comprendre pourquoi débiter l'apprentissage de la lecture dans une des LN des enfants peut faciliter cet apprentissage en français. Lorsqu'ils doivent apprendre à lire dans leur première langue (L1), la plupart des élèves ont, dans cette langue, un niveau de langage oral qui leur permet de comprendre le langage écrit, ce qui n'est pas le cas de la plupart des élèves qui apprennent à lire en français au Sénégal. En effet, d'une part, les enfants maîtrisent mieux leur L1 à l'oral (en particulier le vocabulaire de cette langue) que le français. D'autre part, l'orthographe de ces LN (wolof, pulaar et seereer, en l'occurrence dans le cadre du programme Lecture Pour Tous) est plus régulière que celle du français (parce que leur transcription avec l'alphabet latin est récente), ce qui doit faciliter la compréhension des CGPh et le transfert à une orthographe moins régulière.

Les principaux résultats de l'étude sont les suivants :

- Concernant le **niveau de vocabulaire oral des élèves en français** :
 - A l'entrée à l'école en CI, au début du CP et au début du CEI, le niveau de vocabulaire oral en français des élèves (scores standardisés avec une moyenne de 100 et un écart-type de 15) est respectivement de 89, 102 et 109. Toutefois, en début de CEI, 85% des élèves n'arrivent encore pas à répondre à une seule question orale comportant plusieurs mots français.
 - Le niveau de vocabulaire oral en français augmente donc avec le niveau de scolarité des élèves et le moment où cette augmentation est la plus forte se situe en cours de CI : la différence est de 13 points entre le CI et le CP (presque un écart-type) et de 7 points entre le CP et le CEI.
 - Le niveau de vocabulaire oral en français est :
 - Plus fort parmi les élèves vivant dans un milieu socioculturel favorisé (i.e. niveau d'instruction des parents, maîtrise du français par les parents, accès à des activités préscolaires...)
 - Plus fort lorsque les élèves bénéficient chez eux d'activités dans cette langue (lire, parler et leur apprendre des mots français)
 - Peu dépendant des pratiques actuelles des enseignants sur le français oral
- Concernant le **décodage**
 - Le niveau de décodage en français (évalué par la lecture de mots inventés) augmente avec le niveau de scolarité des élèves et le moment où cette augmentation est la plus forte se situe au cours du CI : augmentation de 86 à 102 entre le CI et le CP et à 112 au CEI (scores standardisés avec une moyenne de 100 et un écart-type de 15). La différence est de 16 points (un peu plus d'un écart-type) entre le CI et le CP et de 10 points entre le CP et le CEI.
- Concernant le **vocabulaire** et la **procédure lexicale d'identification des mots écrits**

- C'est au début de CEI que la grande majorité des élèves (93%) a atteint un niveau de vocabulaire oral en français suffisant pour que cela se traduise de façon positive sur leur niveau de la procédure lexicale d'identification des mots écrits dans cette langue.

Les résultats de cette étude nous amènent à exprimer les **recommandations** suivantes :

1. Les analyses permettent de **cibler le CI comme étant le niveau le plus propice pour travailler intensivement sur le vocabulaire oral en français**. Les pratiques dans ce domaine ne devraient pas se limiter aux mots isolés et devraient tenir compte des différences culturelles (par exemple, travail sur les mots composés et les expressions idiosyncrasiques, spécifiques au français).
2. Le travail sur le **vocabulaire doit se prolonger au CP afin que le niveau de vocabulaire oral en français à l'entrée du CEI permette de commencer l'apprentissage de la lecture dans cette langue dans de bonnes conditions**.
3. **En début de CEI, l'apprentissage de la lecture en français doit cibler en premier sur le décodage et l'identification des mots écrits**, le travail sur la compréhension au-delà du mot s'effectuant principalement à l'oral.
4. Dans la mesure où **les pratiques actuelles des enseignants pour enseigner le français oral n'ont eu que peu d'effets sur l'acquisition du vocabulaire en français** (à la différence de certaines pratiques familiales), **il faudrait revoir les activités conçues dans le curriculum et le matériel didactique** (guides de l'enseignant, récits à lire à haute voix, etc.) pour bien travailler, de façon directe et explicite, la compréhension du français oral (du mot au texte en passant par la phrase).
5. Il en est de même pour les **pratiques autour du décodage, qui doivent s'appuyer sur les caractéristiques de l'orthographe de la langue dans laquelle s'effectue l'apprentissage**.
6. Concernant les **autres activités liées au code, il faudrait réduire le temps que les enseignants disent consacrer à celles sur le nom des lettres**. Cette connaissance peut en effet induire les enfants en erreur : « lm » peut se lire « elle aime », etc.
7. Il faudrait **former les enseignants** pour qu'ils soient en mesure de développer des activités efficaces dans les deux domaines clés pour l'apprentissage de la lecture : décodage et vocabulaire. **Pour faciliter le transfert des acquis dans la lecture en L1 à celle en L2, il faudrait en particulier leur fournir des informations précises sur les similitudes et les différences entre le français et la L1 qu'ils doivent enseigner** au niveau du **décodage** (relations graphèmes-phonèmes) **ainsi que dans les autres domaines** (vocabulaire, morphosyntaxe, entre autres).

Au moment de l'administration de l'enquête sur laquelle cette étude se base, le programme Lecture Pour Tous était déjà mis en œuvre dans certaines classes de CI lors de l'année scolaire

précédente². Il a donc fallu tenir compte du fait que certains élèves ont bénéficié d'heures d'enseignement de la lecture en pular, seereer ou wolof lors de leur année de CI. Cela concerne 310 élèves sur les 399 qui ont été évalués en début de CP. Les analyses menées dans cette étude ont montré que **les élèves ayant** bénéficié d'heures d'enseignement du décodage en LI en CI dans le cadre de Lecture Pour Tous **ont en moyenne un meilleur niveau de vocabulaire oral en français et un meilleur niveau de décodage** que ceux qui n'en ont pas bénéficié (sans qu'on puisse en dégager des éléments d'explication dans le cadre de cette étude). Pour que ce constat soit mieux examiné il serait recommandé de poursuivre les évaluations du niveau de décodage des élèves selon qu'ils aient ou non bénéficié du programme, au-delà du CI (un suivi des élèves sur deux ou trois années pourrait utilement être mené). L'examen des différences entre décodage et procédure lexicale d'identification des mots écrits (qui n'a pas pu être pris en compte au-delà du CI dans cette étude) permettrait notamment d'apporter des éléments d'explication de l'effet surprenant de ce programme (centré sur les langues nationales) sur le niveau de vocabulaire oral en français.

². La construction de l'échantillon de l'enquête sur laquelle cette étude se base a bien entendu tenu compte de cette situation.

INTRODUCTION

Pour améliorer la qualité des apprentissages des élèves, le Gouvernement du Sénégal s'est engagé dans une réforme de l'enseignement de la lecture dans les premières années du cycle primaire. Cet effort bénéficie d'un appui de l'USAID à travers le programme Lecture Pour Tous (2016-2021). De façon plus spécifique, ce programme soutient l'introduction de nouvelles méthodes d'enseignement-apprentissage de la lecture basées sur des preuves scientifiques et l'utilisation pour cet enseignement des langues nationales. Pour le moment, les interventions couvrent le pulaar, le seereer et le wolof pour les élèves des trois premières classes (CI, CP et CEI) des écoles élémentaires publiques et des daaras de sept régions du pays (Diourbel, Fatick, Kaffrine, Kaolack, Louga, Matam, Saint-Louis³).

Une question cruciale dans le cadre du programme est de déterminer le moment le plus approprié pour introduire un enseignement de la lecture en français langue seconde (L2) chez des enfants ayant comme langue première (L1) une des trois langues nationales (LN) suivantes : le pulaar, le seereer ou le wolof. Pour aider le Ministère de l'Education Nationale (MEN) à finaliser le modèle bilingue d'apprentissage-enseignement de la lecture, notamment à fixer le moment le plus approprié pour introduire cet enseignement en français langue seconde (L2), le programme Lecture Pour Tous (qui comprend un volet visant à mener des recherches opérationnelles) a été chargé d'une évaluation du niveau de vocabulaire oral en français des élèves des trois premières classes du cycle primaire.

La présente étude vise une double perspective : (i) une perspective de connaissance du niveau de maîtrise du vocabulaire oral en français des élèves en début de primaire et dans les premières années du cycle ; (ii) une perspective opérationnelle, pour définir le moment et les modalités d'introduction du français, d'un point de vue tant quantitatif (volume de temps dédié) que qualitatif (contenus des activités), et ainsi appuyer le Ministère de l'Education Nationale (MEN) pour finaliser le modèle bilingue de l'enseignement de la lecture qu'il est actuellement en train de définir.

Pour répondre à ces objectifs, nous avons construit et administré des épreuves permettant de mesurer le niveau de vocabulaire oral en français langue seconde (L2) des enfants scolarisés en début CI, CP et CEI, ainsi que leur niveau de prélecture (en début CI et CP) et de lecture (en début CEI). Ce protocole a été complété par trois questionnaires (un destiné aux élèves, un destiné à leurs parents et un destiné aux enseignants que les élèves de CP et de CEI avaient l'année qui a précédé l'enquête), qui visaient à recueillir des informations sur les facteurs linguistiques, sociologiques et pédagogiques susceptibles d'influencer les résultats des élèves à ces évaluations.

Dans la suite de ce document, après avoir présenté rapidement le cadre théorique de l'étude, nous précisons quelle a été sa méthodologie et examinons les principaux résultats en tenant compte de leur évolution en fonction du niveau scolaire, ainsi que de l'incidence exercée sur eux par des facteurs linguistiques, sociologiques et pédagogiques. En conclusion, nous apportons quelques éléments de réponse à la question posée, celle du moment le plus propice

³. La région de Saint-Louis bénéficie seulement de l'appui technique limité du programme Lecture Pour Tous, tandis que les autres bénéficient d'un appui technique compréhensif et un appui financier.

pour introduire un enseignement de la lecture en français L2, ainsi que quelques autres recommandations pour les politiques éducatives, tout en soulignant certaines limites de l'étude

I. APERÇU DU CADRE THEORIQUE

Le cadre théorique de l'étude est présenté en détail dans un document qui accompagne le rapport⁴. Le résumé de ce cadre présenté ci-dessous a pour fonction d'explicitier les choix méthodologiques qui ont été faits, ainsi que nos hypothèses.

Quand ils doivent apprendre à lire, la plupart des élèves ont, dans leur L1, un niveau de langage oral qui leur permet de comprendre les textes écrits. De fait, les recherches signalent que les connaissances précoces de la langue orale (en particulier, le niveau de vocabulaire)⁵ sont de robustes prédicteurs du futur niveau de compréhension en lecture. Toutefois, alors que la compréhension de la langue écrite dépend surtout du niveau de la langue orale chez un bon lecteur, chez un débutant, elle dépend surtout de l'automatisation de la procédure d'identification des mots écrits, qui elle-même, dépend à la fois du niveau du décodage et de celui du vocabulaire oral. Ces résultats ont été relevés dans les apprentissages non seulement en L1 mais aussi en L2 (entre autres, en anglais chez des hispanophones des USA et en français chez des anglophones du Canada)⁶. Quand la procédure d'identification des mots écrits est automatisée, le lecteur peut consacrer ses ressources cognitives à la compréhension de ce qu'il lit, à condition toutefois que son niveau de compréhension de la langue orale soit suffisant, ce qui pose souvent problème quand l'apprentissage s'effectue dans une L2⁷.

Deux autres résultats des recherches sont cruciaux pour la présente étude. Le premier est que la rapidité de l'automatisation du décodage dépend, dans une écriture alphabétique, du degré de régularité des relations entre les plus petites unités du langage écrit, les graphèmes, et les unités correspondantes du langage oral, les phonèmes. En effet, les élèves apprennent plus vite et mieux à lire en espagnol qu'en français, et en français qu'en anglais, ce qui provient du fait que les correspondances graphème-phonème sont plus régulières en espagnol qu'en anglais, l'orthographe du français occupant une position intermédiaire, sachant qu'elle est plus proche de celle de l'espagnol, au moins pour la lecture⁸. Le second résultat important pour le présent projet est que la maîtrise des correspondances graphème-phonème (du décodage) est un puissant mécanisme d'auto-apprentissage⁹, ce qui provient du fait que, quand les correspondances très régulières ont été acquises, il devient plus facile d'apprendre celles qui le sont moins. Ces deux résultats permettent de comprendre pourquoi débiter l'apprentissage de la lecture dans une des LN des enfants peut faciliter cet apprentissage en français. En effet,

⁴. Voir le « Cadre de référence » de ce projet, en partie adapté de différents documents écrits par, sous la direction, ou avec, L. Sprenger-Charolles entre autres : [1] Sprenger-Charolles (2008). Rapport sur l'étude pilote EGRA Sénégal, Banque Mondiale ; [2] RTI (2006). Adaptation en français par L. Sprenger-Charolles du manuel EGRA (seconde édition en 2016) ; [3] Sprenger-Charolles et al. (2019). Présentation du dispositif EvalAide, Ministère de l'éducation nationale et de la jeunesse, France (Voir aussi le chapitre 2 in Dehaene, 2019, pp. 85-158).

⁵. Perfetti (2007) ; Perfetti & Stafura (2014); Quinn, Wagner, Petscher et al. (2015).

⁶. Pour une synthèse sur l'apprentissage en L1, voir Castles et al. (2018) ; en L2, voir August & Shanahan (2006) et Murphy (2018).

⁷. Voir la note précédente.

⁸. Pour une synthèse, voir Ziegler (2016).

⁹. Pour une synthèse, voir Share (1995).

d'une part, l'orthographe de ces LN (wolof, pular et seereer, dans la présente étude) est plus régulière que celle du français¹⁰, ce qui facilite la compréhension des relations graphème-phonème et le transfert à une orthographe moins régulière. D'autre part, les enfants maîtrisent mieux leur LI à l'oral que le français.

En conséquence, il faut aider les enfants à développer des capacités de décodage précises et rapides, ce qui, dans une écriture alphabétique, nécessite un enseignement des correspondances graphème-phonème qui doit être précoce, intensif et systématique. Ce type d'enseignement a un fort impact sur les performances ultérieures des enfants, y compris en compréhension de textes écrits, cet enseignement étant particulièrement bénéfique aux enfants à risque pour des raisons sociologiques ou linguistiques¹¹.

Quand les enfants ont automatisé la procédure d'identification des mots écrits, la compréhension de l'écrit va dépendre essentiellement, comme chez l'adulte, de la capacité de construire une représentation cohérente du message lu, équivalente à celle du message entendu. Plus précisément, pour comprendre un texte, il est nécessaire de mettre en relation plusieurs phrases¹² et, pour cela, il faut souvent effectuer des inférences¹³, c'est-à-dire déduire des informations absentes dans le texte lu (ou entendu). Cette capacité dépend des connaissances que les sujets ont de la langue (LI et L2) et de leurs connaissances du monde, ces dernières étant liées à leur univers culturel. Dans un contexte bilingue, il faut donc prêter attention non seulement aux différences linguistiques qui peuvent gêner la compréhension, mais également aux différences culturelles¹⁴.

Sur la base de ce cadre de référence et des objectifs de cette étude, nous avons élaboré, d'une part, plusieurs épreuves permettant d'évaluer le niveau de vocabulaire oral des élèves en français. Une des épreuves du protocole a été adaptée d'un test souvent utilisé pour évaluer ce niveau chez des enfants de 3 à 8 ans : la désignation de l'image qui, parmi plusieurs, correspond à un mot donné à l'oral par la personne qui fait passer l'épreuve. Les mots sélectionnés (des noms et des verbes), plus ou moins fréquents, sont adaptés à l'âge des élèves examinés. Dans une autre épreuve, les élèves devaient effectuer des actions simples (montrer une partie de leur corps : le « nez », les « yeux », puis le « cou » qui est un terme moins fréquent) ou plus complexes (« mettre un caillou devant [ou derrière] lui »)¹⁵.

D'autre part, nous avons développé une épreuve de prélecture pour le CI et le CP, ainsi qu'une épreuve de lecture pour le CEI. Le choix de deux épreuves différentes se justifie par les résultats des recherches. Les études avec des élèves ayant le français en tant que LI ont en

¹⁰. Parce que leur écriture dans un système alphabétique utilisant les lettres de l'alphabet latin, est récente.

¹¹. Pour des synthèses voir : Early Literacy Panel (2008); National curriculum in England (2013); Rose (2006) ainsi que l'article de Kolinsky et al. (2018) qui signale que les « phonics » méthodes, même si leur efficacité est meilleure chez l'enfant qui apprend à lire quand elles sont utilisées de façon précoce, sont également celles qui s'avèrent être les plus efficaces pour cet apprentissage avec des adultes analphabètes. Pour des données en français, voir Riou & Fontanieu (2016).

¹². Pour une synthèse, voir les parties C3 et C4 (compréhension d'énoncés et de textes) du cadre de référence ; voir aussi Charolles (1995).

¹³. Voir sur les relations entre vocabulaire et inférences, Currie & Cain (2015).

¹⁴. Voir August & Shanahan (2006) et Murphy (2018).

¹⁵. Voir la partie CI (vocabulaire) du cadre de référence. Il est à noter que nous avons demandé de pouvoir faire passer au moins un des tests de vocabulaire non seulement en français mais aussi dans la LN qui est le plus souvent utilisée au Sénégal (le wolof). Cela n'a pas été possible, ce qui limite la portée du présent projet.

effet montré que les scores en lecture de mots isolés augmentent progressivement entre le début et le milieu du CP et, surtout, entre le milieu du CP et la fin de cette classe : l'évolution étant alors tant quantitative que qualitative, comme l'indiquent certains changements observés entre le milieu et la fin du CP pour la lecture de différents types de mots. En effet, en milieu de CP, les mots réguliers fréquents (comme « table ») ne sont pas mieux lus que les mots inventés (comme « tople »), et les scores sur les mots irréguliers même très fréquents sont très bas, ces items donnant lieu à de nombreuses erreurs phonologiques (« sept » lu comme « septembre »). Ces résultats signalent que les élèves utilisent alors essentiellement le décodage, mais pas encore la procédure lexicale d'identification des mots écrits. Quelques mois plus tard (fin CP), le tableau est différent. Les mots réguliers sont alors mieux lus que les mots inventés¹⁶. L'avantage des mots réguliers sur ceux qui sont inventés (qui n'a pas été relevé dans une étude avec des enfants sénégalais dans les trois premières classes)¹⁷ est le premier signe d'un recours à la procédure lexicale d'identification des mots écrits. Il s'explique par le fait que les mots réguliers bénéficient de la régularité et de la fréquence, ce qui n'est pas le cas pour les mots inventés, qui ne bénéficient que de la régularité. En revanche, à la même époque les scores sur les mots irréguliers (qui ne bénéficient que de la fréquence) sont encore très faibles¹⁸.

Enfin, trois questionnaires, qui visaient à recueillir des informations sur les facteurs linguistiques, sociologiques et pédagogiques susceptibles d'influencer les résultats des évaluations, ont été développés : un était destiné aux élèves, un autre à leurs parents et un dernier aux enseignants que les élèves actuellement en CP et en CEI avaient l'année précédente (respectivement en CI et en CP).

Principales hypothèses de la présente étude

Quand ils doivent apprendre à lire dans leur L1, la plupart des élèves ont, dans cette langue, un niveau de langage oral qui leur permet de comprendre le langage écrit. Ce qui va surtout poser problème à un lecteur débutant dans une L2, est lié au fait que son niveau de compréhension orale dans cette L2, et plus particulièrement son niveau de vocabulaire¹⁹, n'est souvent pas suffisant pour lui permettre de comprendre ce qu'il lit dans cette langue. En conséquence :

- La première hypothèse (H1) est que pour pouvoir comprendre ce qu'ils lisent en français, les élèves devraient non seulement être capables de comprendre à l'oral des mots isolés, mais aussi des suites de quelques mots.
- La deuxième hypothèse (H2) comprend trois volets, dont les deux premiers sont reliés à la première hypothèse : le niveau de vocabulaire oral en français de ces élèves (H2^{voc}), tout comme leur niveau de prélecture ou de lecture (H2^{lec}), devraient augmenter avec leur niveau de scolarité. Cette évolution devrait permettre de cerner à partir de quel degré de maîtrise du français oral (mots isolés et groupes de mots) il devient possible d'apprendre à lire dans cette L2 (H2^{voc-lec}).

¹⁶. Voir l'étude longitudinale de Sprenger-Charolles et al. (2003), ainsi que la synthèse des études francophones par Deacon et al. (2017)

¹⁷. Sprenger-Charolles (2008). Rapport sur l'étude pilote EGRA Sénégal, Banque Mondiale.

¹⁸. Voir l'étude longitudinale de Sprenger-Charolles et al. (2003) et la synthèse des études francophones par Deacon et al. (2017); voir aussi la partie B1.2 du cadre de référence (Évaluation des procédures d'identification des mots).

¹⁹. Voir August & Shanahan (2006) et Murphy (2018).

Toutefois, pour qu'il comprenne ce qu'il lit dans une écriture alphabétique, un élève doit non seulement avoir un certain niveau de compréhension de la langue orale dans laquelle il lit, en L1 comme en L2, il lui faut également avoir automatisé le décodage. Comme nous l'avons déjà signalé, la rapidité de cette automatisation dépend du degré de régularité des relations entre les plus petites unités du langage écrit, les graphèmes, et les unités correspondantes du langage oral, les phonèmes. De plus, la maîtrise du décodage (des correspondances graphème-phonème) est un puissant mécanisme d'auto-apprentissage, ce qui provient du fait que, quand les correspondances les plus régulières ont été acquises, il devient plus facile d'apprendre celles qui le sont moins.

Ces données²⁰ permettent de comprendre pourquoi débiter l'apprentissage de la lecture dans une des LN des enfants peut faciliter cet apprentissage en français. En effet, d'une part, l'orthographe de ces LN (ici, du wolof, du pulaar et du seereer) est plus régulière que celle du français, ce qui facilite la compréhension des relations graphème-phonème et le transfert à une orthographe moins régulière. D'autre part, les enfants maîtrisent mieux leur L1 à l'oral que le français. En particulier, leur niveau de vocabulaire oral dans leur L1 est supérieur à celui qu'ils ont en français L2, ce qui doit avoir une incidence sur la procédure lexicale d'identification des mots écrits français. Les prédictions précédentes peuvent être évaluées, au moins partiellement, par les résultats obtenus en CEI dans l'épreuve de lecture qui, à ce niveau scolaire, contient des items permettant d'examiner le fonctionnement du décodage et de la procédure lexicale d'identification des mots écrits : le décodage (par la lecture de mots inventés) et la procédure lexicale (par la lecture de mots issus du français mais aussi de leur LN). Dans ce cadre, il est possible d'avancer une série d'hypothèses concernant les relations entre le niveau de vocabulaire oral en français et celui du décodage et de la procédure lexicale d'identification des mots écrits.

- La troisième hypothèse (H3) comprend deux volets : si les élèves de CEI ont recours, en plus de la procédure phonologique (le décodage), à la procédure lexicale d'identification des mots écrits, les mots de leur LN devraient être mieux lus que les mots inventés ou les mots en français (H3^{lec}). De plus, leur niveau de vocabulaire oral en français devrait rendre compte de leurs scores en lecture de mots dans cette langue (H3^{voc-lec}).
- La quatrième série d'hypothèses (H4) concerne les activités en français. Alors que celles autour du vocabulaire et de la compréhension orale devraient surtout avoir une incidence positive sur le niveau de vocabulaire oral en français (H4^{voc1}), celles autour du code devraient faciliter l'acquisition du décodage (H4^{lec1}). Il est aussi possible de prédire que la participation au programme Lecture Pour Tous doit avoir une incidence positive sur le décodage (H4^{lec2}) et, en retour, sur la procédure lexicale d'identification des mots écrits (H4^{lec3}), ainsi que sur le niveau de vocabulaire oral dans les LN (H4^{lec-voc}).

Ces différentes hypothèses ne prennent pas en compte les caractéristiques des élèves et de leur environnement familial et scolaire (pour ceux qui sont déjà à l'école). Ces facteurs contextuels pourraient avoir une influence sur le niveau de vocabulaire oral en français des élèves et/ou sur le niveau de la procédure d'identification des mots écrits.

²⁰. Pour la L1, voir la synthèse de Castles et al. (2018) et, en L2, voir August & Shanahan (2006) et Murphy (2018).

- La cinquième série d'hypothèses (H5) est que certaines de ces caractéristiques ont un lien avec l'environnement de l'enfant hors de l'école, en particulier son milieu social (H5¹). Le niveau de vocabulaire oral en français, tout comme celui de la procédure d'identification des mots écrits, sont supposés être meilleurs chez les élèves issus d'un milieu socioculturel favorisé (H5^{voc1} pour le vocabulaire²¹ et H5^{lec1} pour l'identification des mots écrits)²². De plus, les enfants qui bénéficient, dans leur entourage, d'une personne qui passe du temps à leur parler, à leur apprendre des mots nouveaux ou encore à leur lire des histoires en français (H5²), devraient avoir un meilleur niveau de vocabulaire oral (H5^{voc2}) dans cette langue, et un meilleur niveau de lecture (H5^{lec2}), que ceux qui ne bénéficient pas d'un environnement de ce type²³.

- La sixième série d'hypothèses (H6) concerne les facteurs en lien avec l'environnement scolaire de l'enfant, en particulier, la place qui est donnée aux activités d'enseignement-apprentissage du vocabulaire oral en français aux activités sur le décodage (cf. ci-dessus les hypothèses H4^{lec}, H4^{voc} et H4^{lec-voc}). Sont également supposées avoir un effet facilitateur sur l'acquisition du vocabulaire oral en français (H6^{voc}) ainsi que sur les procédures spécifiques à la lecture dans cette langue (décodage et procédure lexicale d'identification des mots écrits, H6^{lec}), les compétences des enseignants en français telles qu'elles peuvent être appréhendées par leur niveau de formation (H6^{voc1} et H6^{lec1}) ou par la perception qu'ils ont de leurs compétences dans cette langue (H6^{voc2} et H6^{lec2}).

²¹. Pour la L1, voir pour des données en français : Gentaz et al. (2013), et, pour des données en anglais : Hoff (2006) ; Noble & McCandliss (2005) ; en L2, voir les synthèses de August & Shanahan (2006) et de Murphy (2018).

²². Si le niveau d'identification des mots écrits est plus bas chez des enfants de milieu socioculturel défavorisé, la différence en fonction du milieu est toutefois plus importante pour ce qui relève du domaine de la compréhension, y compris celle du vocabulaire, et ce surtout en L2, cf. : August & Shanahan (2006) et Murphy (2018).

²³. Dans la mesure où le niveau socioculturel de la famille (qui intègre le niveau de scolarisation des parents, leur savoir lire et leur profession) est fortement corrélé au degré d'aisance en français, aucune hypothèse ne prend en compte uniquement le niveau de maîtrise du français des parents.

2. METHODOLOGIE

Cette section présente les modalités de construction des instruments de l'enquête sur lesquels notre étude s'appuie (2.1), puis les principes qui ont guidé la sélection de l'échantillon, ainsi que les caractéristiques de cet échantillon (2.2) et enfin, le protocole d'évaluation des compétences des élèves (2.3).

2.1. Présentation de l'enquête

Dans la perspective d'évaluer le niveau de vocabulaire oral en français des élèves des premières années du primaire, un protocole visant à examiner leurs compétences dans ce domaine (ainsi qu'en lecture) a été construit et administré à des élèves de CI, de CP et de CEI (Instrument 1a, qui sera présenté dans la section 2.3). Une fois les niveaux de compétences des élèves connus, les analyses consistent alors à chercher à expliquer leur variabilité. Deux pistes d'explication complémentaires sont envisagées.

La première se situe dans la sphère privée des élèves : elle consiste à estimer dans quelle mesure les caractéristiques des élèves (âge, sexe, parcours préscolaire, etc.) et de leur environnement familial (situation socioculturelle, linguistique, proximité avec le français dans leur vie quotidienne, etc.) peuvent expliquer la variabilité des scores qu'ils ont obtenus au test. Pour obtenir ces informations, un certain nombre de questions ont été posées directement aux élèves (Instrument 2) et à un de leurs parents (Instrument 3). Si cette piste d'explication est intéressante en elle-même, les variables qui la composent serviront également de contrôle dans les analyses suivantes (pour raisonner à caractéristiques et environnement équivalents).

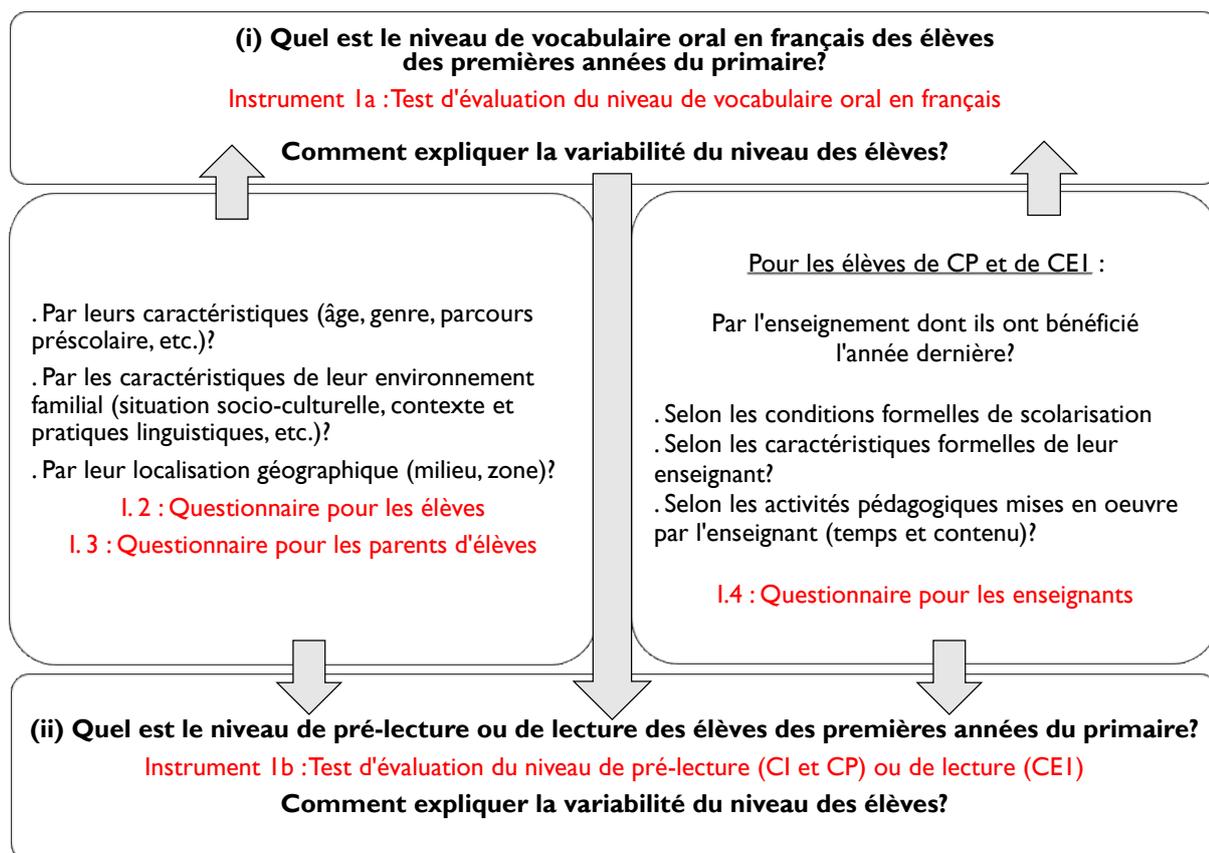
La seconde piste est liée plus directement à la sphère scolaire, c'est à dire à l'enseignement dont les élèves ont bénéficié l'année précédant l'enquête. Plusieurs dimensions sont là aussi mobilisées : (i) les caractéristiques formelles de la classe (le nombre d'élèves, la disponibilité de manuels, etc.) ; (ii) les caractéristiques formelles de l'enseignant (son niveau de formation initiale et continue, son nombre d'années d'expérience, etc.) ; (iii) la perception des enseignants de leur aisance en français ; (iv) le temps alloué à certaines activités d'enseignement-apprentissage du vocabulaire oral en français et (v) le contenu des activités mises en œuvre pour l'enseignement-apprentissage du vocabulaire oral en français. Ces renseignements, qui ne concernent que les enfants de CP (qui étaient en CI l'année précédente) et de CEI (qui étaient en CP l'année précédente), sont fournis par des questionnaires administrés aux enseignants que les élèves avaient l'année précédant l'enquête (Instrument 4).

Pour définir le moment le plus propice pour introduire un enseignement de la lecture en français L2, le niveau de vocabulaire oral en français des élèves doit être mis en regard avec leur niveau de prélecture ou de lecture. Un protocole visant à examiner leurs compétences en prélecture et en lecture a donc été construit et administré aux élèves (Instrument 1b, présenté dans la partie B3). Une fois les niveaux de prélecture et de lecture des élèves connus, la perspective est de chercher à expliquer leur variabilité d'une part, par les caractéristiques des élèves et de leur environnement familial, d'autre part et pour les élèves de CP et de CEI, par l'enseignement dont ils ont bénéficié l'année précédant l'enquête, et enfin, par leur niveau de vocabulaire oral en français. Il s'agit d'analyser si et dans quelle mesure le niveau de vocabulaire oral en français des élèves a un effet sur leur niveau de prélecture ou de lecture, et d'identifier

s'il existe un niveau minimal de vocabulaire qu'il est nécessaire que les élèves aient pour aborder de façon optimale l'apprentissage de la lecture.

Les informations utiles pour mener à bien cette étude ont donc nécessité la construction de différents instruments et leur administration à différents acteurs (graphique 1 ci-dessous). Ces différentes données ont ensuite été consolidées et fusionnées dans une même base, ce qui permet de disposer des résultats de chaque élève à chaque item du protocole d'évaluation, de connaître ses caractéristiques et celles de ses parents, ainsi que, pour ceux de CP et de CEI, d'avoir des informations sur ce qu'ils ont fait en classe l'année précédente.

Graphique 1 : Présentation des instruments de l'étude



2.2. Présentation de l'échantillon

Avant de présenter les critères qui ont guidé la sélection de l'échantillon, il convient de rappeler que le programme Lecture Pour Tous soutient l'introduction de l'enseignement de la lecture en pular, seereer et wolof pour les élèves des trois premières classes du primaire (CI, CP et CEI) dans les écoles élémentaires publiques et les daaras de sept régions du pays (Diourbel, Fatick, Kaffrine, Kaolack, Louga, Matam, Saint-Louis). La mise en œuvre du programme a débuté à la rentrée 2017-2018 (an 1 du programme) pour les élèves de CI de 487 écoles des régions de Fatick, Kaffrine et Kaolack, puis a concerné, en 2018-2019 (an 2 du programme), les élèves de CI de 2 263 écoles des régions de Diourbel, Fatick, Kaffrine, Kaolack, Louga et Matam. Il faut également noter que la langue nationale utilisée pour

l'enseignement de la lecture (pulaar, wolof ou seereer) est déterminée au niveau de chaque école.

2.2.1. Les critères de sélection

De façon logique, l'enquête a ciblé les zones d'intervention du programme. Pour des raisons logistiques, trois régions proches les unes des autres ont été choisies : celles de Fatick, Kaolack et Kaffrine. Parmi elles, trois Inspections d'Education et de Formation (IEF) ont ensuite été sélectionnées de façon aléatoire : l'IEF de Gossas dans l'Inspection d'Académie (IA) de Fatick, l'IEF de Kaolack Commune dans l'IA de Kaolack et l'IEF de Kaffrine dans l'IA de Kaffrine. La construction de l'échantillon de l'enquête a ensuite cherché à respecter (i) la répartition du nombre d'élèves scolarisés entre ces trois IEF, puis au sein de chaque IEF, la répartition (ii) selon le milieu de résidence (urbain / rural) et (iii) selon la langue choisie pour l'école et enfin, (iv) la répartition des écoles selon l'année de mise en place du programme (an 1 ou an 2). Une fois ces critères appliqués, un certain nombre d'écoles ont été sélectionnées selon des modalités aléatoires. Un travail a ensuite consisté à contacter les directeurs d'école pour savoir si les enseignants qui tenaient les classes de CI et de CP en 2017-2018 étaient encore présents dans l'école pour l'année 2018-2019. Quand ce n'était pas le cas, l'école a été remplacée pour une autre comportant les mêmes caractéristiques (IEF, milieu, langue nationale choisie pour l'école, année de mise en place du programme). Enfin, au sein de chaque école élémentaire, le test a été administré à dix élèves (non redoublants et en essayant de respecter un équilibre entre garçons et filles) de CI, de CP et de CEI. Au total, l'échantillon compte 1 183 élèves : 388 de CI, 399 de CP et 396 de CEI. Leur répartition entre les différentes IEF selon le milieu, la langue de l'école et la date d'entrée dans le programme Lecture Pour Tous est présentée dans le tableau I ci-dessous.

Tableau I : Distribution des écoles et des élèves enquêtés par région

IA		Fatick	Kaffrine	Kaolack	Ensemble
IEF		Gossas	Kaffrine	Kaolack Commune	
Nombre d'écoles - d'élèves		15 - 449	21 - 554	6 - 180	42 - 1 183
Milieu	Urbain	1 - 30	1 - 30	6 - 180	8 - 240
	Rural	14 - 419	20 - 524	0 - 0	34 - 943
Langue	Pulaar	1 - 30	0 - 0	0 - 0	1 - 30
	Seereer	5 - 149	1 - 30	1 - 30	7 - 209
	Wolof	9 - 270	20 - 524	5 - 150	34 - 944
Année d'entrée LPT	An 1 (2017-18)	12 - 359	13 - 379	5 - 150	30 - 888
	<i>Dont CP²⁴</i>	<i>12 - 120</i>	<i>14 - 140</i>	<i>4 - 40</i>	<i>30 - 310</i>
	An 2 (2018-19)	3 - 90	7 - 175	1 - 30	11 - 295

2.2.2. Les caractéristiques des élèves de l'échantillon

Les informations concernant les caractéristiques des élèves ont été recueillies directement auprès d'eux, ainsi qu'auprès d'un de leurs parents. Nous examinerons d'abord les

²⁴. Précisons que ces élèves ont bénéficié l'année précédant l'enquête (lorsqu'ils étaient donc en CI) d'activités d'enseignement en pulaar, seereer ou wolof dans le cadre du programme Lecture Pour Tous.

caractéristiques propres aux élèves et à leur environnement familial, puis celles de leurs parents, avant d'examiner le contexte linguistique dans lequel ils vivent.

Les caractéristiques formelles des élèves et de leur famille

Le tableau 2 ci-après nous renseigne sur un certain nombre de caractéristiques des élèves de l'échantillon.

Tableau 2 : Distribution des élèves selon certaines de leurs caractéristiques

	CI		CP		CEI		Ensemble	
	N	%	N	%	N	%	N	%
Nombre d'élèves	388	-	399	-	396	-	1 183	-
Genre								
Garçons	187	48,2%	197	49,4%	184	46,5%	568	48,0%
Filles	201	51,8%	202	50,6%	212	53,5%	615	52,0%
Age (année)								
Moyen	7,2		8,3		9,8		8,4	
Min-Max	5-13		6-18		6-16		5-18	
Précolaire								
Non	333	85,8%	282	70,7%	289	73,0%	904	76,4%
Oui	55	14,2%	113	28,3%	105	26,5%	273	23,1%
Durée de préscolarisation								
1 an	15	27,3%	32	28,3%	29	27,6%	76	27,8%
2 ans	8	14,5%	47	41,6%	41	39,0%	96	35,2%
3 ans	14	25,5%	34	30,1%	35	33,3%	83	30,4%
NSP	18	32,7%	0	0,0%	0	0,0%	18	6,6%
Type de préscolaire								
Classe pré-élémentaire	3	5,5%	0	0,0%	1	1,0%	4	1,5%
Case des Tout Petits	30	54,5%	14	12,4%	3	2,9%	47	17,2%
Maternelle privée	0	0,0%	1	0,9%	4	3,8%	5	1,8%
Maternelle franco-arabe	1	1,8%	27	23,9%	15	14,3%	43	15,8%
Communautaire	0	0,0%	10	8,8%	11	10,5%	21	7,7%
Daara	3	5,5%	61	54,0%	71	67,6%	135	49,5%
Vit avec ses deux parents								
Oui	326	84,0%	316	79,2%	322	81,3%	964	81,5%
Non, avec un seul	36	9,3%	42	10,5%	39	9,8%	117	9,9%
Non, avec aucun des deux	24	6,2%	38	9,5%	35	8,8%	97	8,2%
Fratrie								
Non	15	3,9%	16	4,0%	14	3,5%	45	3,8%
Oui, plus jeunes	72	18,6%	79	19,8%	80	20,2%	231	19,5%
Oui, plus âgés	55	14,2%	47	11,8%	40	10,1%	142	12,0%
Oui, plus jeunes et plus âgés	244	62,9%	254	63,7%	262	66,2%	760	64,2%

Dans le tableau 2 ci-dessus, une première information concerne le genre des élèves : les filles sont un peu plus nombreuses que les garçons (52% contre 48%). S'agissant de l'âge des élèves, la moyenne s'établit à 7 ans au CI, 8 ans au CP et 10 ans en CEI, sachant que pour chacune de ces classes, il varie de 5 à 13 ans parmi les élèves de CI, de 6 à 18 ans parmi ceux de CP et de 6 à 16 ans parmi ceux de CEI (les extrêmes étant bien sûr peu fréquents). Enfin, la grande majorité des élèves (un peu plus de 80%) vit avec ses parents ; seuls 10% ne vit qu'avec un seul de ses parents et 8% avec aucun des deux. Par ailleurs, quasiment tous les élèves (96%) ont des frères et sœurs.

En ce qui concerne le parcours préscolaire des élèves, il faut d'abord noter que les informations concernant les élèves de CI n'ont pas été collectées au moment de l'enquête, mais quelques semaines après, à partir d'une fiche remise aux directeurs d'école. Il semblerait que les consignes pour le remplissage des fiches n'aient pas été suffisamment explicitées. On observe en effet que le taux de préscolarisation des élèves de CI est nettement plus faible que celui enregistré pour les élèves de CP et de CEI (tableau 2 ci-dessus: respectivement, 14% contre 28% et 26%). Le nombre d'élèves ayant fréquenté un daara préscolaire étant beaucoup plus faible parmi les élèves de CI (5%) que parmi ceux des deux autres classes (54 et 68%), on peut émettre l'hypothèse que certains élèves ayant fréquenté ce type de préscolaire n'ont pas été considérés comme ayant été préscolarisés. Il faut donc considérer ces données avec précaution. Parmi les élèves de CP et de CEI qui ont été préscolarisés, il existe une certaine variété dans la durée de préscolarisation, allant d'une seule année à trois ans (une courte majorité d'enfants l'ayant été pendant deux ans).

Pour recueillir des informations sur les caractéristiques propres aux parents ou personnes en charge des élèves, l'enquêteur a interrogé un parent, qui a indiqué, pour lui et l'autre parent²⁵, leur niveau de savoir lire et leur niveau d'instruction. Ces informations sont présentées dans le tableau 3 ci-après. Il nous a semblé plus pertinent de considérer de façon combinée les renseignements des deux parents sur leur savoir lire et leur niveau d'instruction, pour essayer de mieux appréhender le contexte de l'enfant sur ces deux dimensions.

Tableau 3 : Distribution des élèves selon le savoir lire et le niveau d'instruction des parents

	CI		CP		CEI		Ensemble		
	N	%	N	%	N	%	N	%	
Nombre d'élèves	388		399		396		1 183		
Savoir lire des parents									
Aucun ne sait lire	161	41,5%	134	33,6%	142	35,9%	437	36,9%	
Un seul sait lire	139	35,8%	165	41,4%	163	41,2%	467	39,5%	
Les deux savent lire	88	22,7%	100	25,1%	91	23,0%	279	23,6%	
Niveau d'instruction des parents									
	Parent 1	Parent 2							
Non scolarisé	Non scolarisé	194	50,0%	176	44,1%	182	46,0%	552	46,7%
Non scolarisé	Alpha / primaire	98	25,3%	107	26,8%	109	27,5%	314	26,5%
Non scolarisé	Secondaire ou +	30	7,7%	33	8,3%	32	8,1%	95	8,0%
Alpha / primaire	Alpha / primaire	28	7,2%	39	9,8%	44	11,1%	111	9,4%
Alpha et +	Secondaire et +	38	9,8%	44	11,0%	29	7,3%	111	9,4%

²⁵. Pour 16 enfants de l'échantillon, un seul parent est concerné.

Si un peu plus d'un tiers des enfants (37%) ont des parents qui ne savent pas lire, ils sont en revanche 39% à avoir un de leurs parents alphabète et 24% à avoir deux parents alphabètes. Cette tendance est peu ou prou la même pour les élèves des trois classes. En outre, les parents de près de la moitié des enfants (47%) n'ont jamais été scolarisés. Le niveau de scolarisation de l'autre moitié des enfants se répartit comme suit : 34% d'entre eux ont un parent qui n'a jamais été scolarisé, l'autre ayant atteint le cycle primaire ou participé à un programme d'alphabétisation (26%), soit l'enseignement secondaire ou plus (8%). Enfin, 19% des enfants ont deux parents qui ont été scolarisés (9% ont des parents qui ont fréquenté un programme d'alphabétisation ou atteint le cycle primaire, et 9% ont un parent dans cette même situation et un autre qui a atteint l'enseignement secondaire ou plus).

Des questions ont également été posées aux élèves pour appréhender la place de la lecture chez eux. Plusieurs dimensions sont ainsi examinées : la présence de livres à la maison et leur utilisation par les élèves, la présence d'une personne sachant lire à la maison (sans que ce soit forcément un parent) et si les élèves passent du temps à regarder ou lire des livres avec cette personne. Ces informations sont présentées dans le tableau 4 ci-après.

Tableau 4 : Distribution des élèves selon leur relation à la lecture chez eux

	CI		CP		CEI		Total	
	N	%	N	%	N	%	N	%
Livres à la maison	60	15,5%	190	47,6%	213	53,8%	463	39,1%
Regarde / lit	52	13,4%	182	45,6%	201	50,8%	435	36,6%
Ne regarde / lit pas	8	2,1%	8	2,0%	12	3,0%	28	2,4%
Pas de livre à la maison	328	84,5%	209	52,4%	183	46,2%	720	60,9%
Ensemble	388	100,0%	399	100,0%	396	100,0%	1183	100,0%
Quelqu'un sait lire à la maison	243	62,6%	314	78,7%	338	85,4%	895	75,7%
Regarde / lit avec	159	41,0%	186	46,6%	217	54,8%	487	47,5%
Ne regarde / lit pas avec	84	21,6%	128	32,1%	121	30,6%	408	28,1%
Personne ne sait lire	145	37,4%	85	21,3%	58	14,6%	288	24,3%
Ensemble	388	100,0%	399	100,0%	396	100,0%	1183	100,0%

Un peu plus des deux tiers des élèves (39%) déclarent qu'il y a des livres chez eux et parmi eux, la grande majorité (94%) dit les regarder ou les lire. Au-delà de cette tendance générale, on observe que ces chiffres augmentent en fait avec la classe de l'élève. En effet, la proportion d'élèves déclarant avoir des livres chez eux passe de 15% parmi les élèves de CI à 48% parmi ceux de CP et à 54% parmi ceux de CEI. Au total, 13% des élèves de CI déclarent regarder ou lire des livres chez eux, alors qu'ils 46% parmi les élèves de CP et 51% parmi ceux de CEI.

La grande majorité des élèves (76%) dit qu'une personne de leur entourage proche (sans que cela soit forcément un parent) sait lire et un peu plus de la moitié (54%) déclare que cette personne passe du temps à regarder ou lire des livres avec eux. On trouve là encore que la proportion d'élèves indiquant passer du temps avec quelqu'un à regarder ou lire des livres augmente avec la classe. En effet, si de façon globale, ce chiffre s'élève à 47%, il varie en fait de 41% parmi les élèves de CI à 47% parmi ceux de CP et à 55% parmi ceux de CEI.

Les informations précédentes ont été combinées pour savoir dans quelle mesure les enfants passent du temps à regarder ou lire des livres et comment ils le font (tableau 5 ci-après). Seul un quart des élèves de CI passe du temps à regarder ou lire des livres chez eux, 9% le faisant

seul et avec quelqu'un qui sait lire. Cette proportion s'élève à 60% parmi les élèves de CP et à 70% avec ceux de CEI ; parmi eux, un tiers le fait seul et avec quelqu'un qui sait lire.

Tableau 5 : Distribution des élèves selon leur pratique générale de lecture

	CI		CP		CEI		Total	
	N	%	N	%	N	%	N	%
Ne lit pas	286	73,7%	162	40,6%	120	30,3%	568	48,0%
Lit seul ou avec quelqu'un	67	17,3%	102	25,6%	131	33,1%	300	25,4%
Lit seul et avec quelqu'un	35	9,0%	135	33,8%	145	36,6%	315	26,6%
Ensemble	388	100,0%	399	100,0%	396	100,0%	1183	100,0%

En résumé, une minorité d'élèves a bénéficié d'activités préscolaires (14% des élèves de CI, 28% de ceux de CP et 26% de ceux de CEI). Les enfants des deux dernières classes ont surtout fréquenté un daara (54% pour ceux de CP, 68% pour ceux de CEI, contre environ 5% pour ceux de CI). Les différences relevées entre les élèves de CI et ceux des deux autres classes sont toutefois à prendre avec précaution : certains élèves de CI ayant fréquenté un daara ont sans doute été considérés comme n'ayant pas été préscolarisés. Concernant l'environnement familial des élèves, environ un élève sur deux a au moins un parent qui a été scolarisé, la majorité ayant atteint au mieux le niveau primaire. De même, la majorité d'entre eux (de 59 à 66% selon la classe) a au moins un parent sachant lire. Enfin, la moitié des élèves indique pratiquer la lecture à la maison et disposer de livres, ces proportions augmentant avec le niveau scolaire : de 26% en CI à 59% en CP et à 70% en CEI pour les activités de lecture ; de 25% en CI à 48% en CP et 54% en CEI, pour la présence de livres à la maison.

Le contexte linguistique des élèves

Une série d'informations collectées tant auprès des élèves que de leurs parents concerne l'environnement linguistique, dans lequel les élèves vivent. Cette dimension est importante : les langues que les enfants utilisent dans leur vie quotidienne peuvent différer sur de nombreux domaines du langage, ce qui peut avoir une incidence sur l'apprentissage de la lecture. Ces domaines concernent :

- **La phonologie** : par exemple, différences dans le nombre et la nature des consonnes et des voyelles (cf. B4.1 dans le cadre de référence) ;
- **L'orthographe** : par exemple, différences de notation des voyelles écrites dans les langues qui ont plus de six voyelles (rappel : il n'y a que 6 lettres dans l'alphabet pour les écrire, cf. B4.2 dans le cadre de référence) ;
- **La morphologie grammaticale** : par exemple, certaines langues ont un genre neutre alors que d'autres ne marquent pas le genre ; autre exemple, le marquage du passé et du futur, qui peut être intégré au verbe (comme en français : *il est venu, il viendra, il va venir*) ou être uniquement porté par des mots comme *hier, demain, avant, après...*, sans aucune marque sur le verbe (cf. C2.1 et C2.2 dans le cadre de référence) ;
- **La morphologie lexicale** : par exemple, présence de classifieurs dans des langues d'Afrique, comme <ier> qui désigne le nom d'un métier ou d'un arbre en français (cf. C2.2).

D'autres différences sont plus liées à la culture : par exemple, de très nombreux termes peuvent désigner les différentes façons de marcher dans certaines langues d'Afrique (cf. C1.1

dans le cadre de référence) ; d'autres exemples sont présentés dans la partie concernant les inférences basées sur les connaissances du monde (cf. C4.2 dans le cadre de référence).

Deux dimensions sont examinées dans les questionnaires : la ou les langues que les élèves utilisent dans leur vie quotidienne et le degré de familiarité des élèves avec le français. Avant cela, la première question concerne, de façon classique, la langue maternelle des parents d'élèves. Les informations sont présentées dans le tableau 6 ci-dessous.

Tableau 6 : Distribution des élèves selon la langue maternelle de leurs parents

Langue maternelle des parents	CI		CP		CEI		Total	
	N	%	N	%	N	%	N	%
Une seule langue maternelle	355	91,5%	363	91,0%	367	92,7%	1 085	91,7%
Wolof	215	55,4%	231	57,9%	217	54,8%	663	56,0%
Pulaar	38	9,8%	38	9,5%	48	12,1%	124	10,5%
Seereer	93	24,0%	90	22,6%	85	21,5%	268	22,7%
Autre	9	2,3%	4	1,0%	17	4,3%	30	2,5%
Deux langues maternelles	33	8,5%	36	9,0%	29	7,3%	98	8,3%
Total	388	100,0%	399	100,0%	396	100,0%	1 183	100,0%

La très grande majorité des enfant (92%) vit dans une famille où les parents ont la même langue maternelle : le wolof dans 56% des cas, le seereer dans 23%, le pulaar dans 10% et une autre langue que celles-ci dans 2% des cas. Il n'y a que dans 8% des familles que les parents ont des langues maternelles différentes. Cette tendance se retrouve dans les trois classes.

La langue maternelle ne suffit cependant pas à caractériser l'environnement linguistique familial, les parents pouvant ne pas utiliser leur langue maternelle pour parler à l'enfant. Il est donc utile d'examiner la ou les langues effectivement parlées à la maison (tableau 7 ci-dessous). La très grande majorité (83%) des élèves parle une seule langue à la maison. Il s'agit du wolof pour 64% d'entre eux, du seereer pour 14%, du pulaar pour 5% et d'une autre langue pour 1%. On trouve ainsi 17% d'élèves, qui parlent au moins deux langues chez eux.

Tableau 7 : Distribution des élèves selon la langue parlée à la maison

Langue parlée à la maison	CI		CP		CEI		Total	
	N	%	N	%	N	%	N	%
Une seule langue	317	81,7%	336	84,2%	331	83,6%	984	83,2%
Wolof	244	62,9%	264	66,2%	248	62,6%	756	63,9%
Pulaar	17	4,4%	20	5,0%	19	4,8%	56	4,7%
Seereer	55	14,2%	52	13,0%	57	14,4%	164	13,9%
Autre	1	0,3%	0	0,0%	7	1,8%	8	0,7%
Deux langues ou plus	71	18,3%	63	15,8%	65	16,4%	199	16,8%
Total	388	100,0%	399	100,0%	396	100,0%	1 183	100,0%

Les élèves ont également été interrogés sur la ou les langues qu'ils utilisent avec leurs amis. Ces informations sont données dans le tableau 8 ci-après. Il ressort que la grande majorité des élèves parle une seule langue avec leurs amis : le wolof pour 71% d'entre eux, le seereer pour 14% et le pulaar pour 2%. Ils sont 12% à déclarer parler deux langues avec leurs amis (dont 1% à dire qu'ils parlent français) et 1% à utiliser trois langues.

Tableau 8 : Distribution des élèves selon la langue parlée avec leurs amis

Langue parlée avec les amis	CI		CP		CEI		Total	
	N	%	N	%	N	%	N	%
Une langue	341	87,9%	353	88,5%	338	84,1%	1032	86,8%
Wolof	273	70,4%	293	73,4%	274	68,2%	840	70,6%
Pulaar	9	2,3%	8	2,0%	10	2,5%	27	2,3%
Seereer	59	15,2%	52	13,0%	53	13,2%	164	13,8%
Autre	0	0,0%	0	0,0%	1	0,2%	1	0,1%
Deux langues	45	11,6%	42	10,5%	55	13,7%	142	11,9%
Dont le français	3	0,8%	4	1,0%	9	2,2%	16	1,3%
Trois langues	2	0,5%	4	1,0%	9	2,2%	15	1,3%
Ensemble	388	100,0%	399	100,0%	402	100,0%	1189	100,0%

A partir des renseignements fournis par les parents et les élèves sur la ou les langues parlées à la maison et avec les amis, nous avons cherché à identifier combien et quelles sont celles qui sont le plus utilisées par les élèves en dehors de l'école (tableau 9 ci-dessous). En fait, dans leur vie quotidienne, 25% des élèves utilisent au moins deux langues, les 75% restant évoluent dans un contexte monolingue : 62% utilisent uniquement le wolof, 11% le seereer et 2% le pulaar.

Tableau 9 : Distribution des élèves selon le nombre de langues utilisées dans la vie quotidienne

Langues dans la vie quotidienne	CI		CP		CEI		Total	
	N	%	N	%	N	%	N	%
Une langue	294	75,8%	303	75,9%	291	73,5%	888	75,1%
Wolof	242	62,4%	256	64,2%	233	58,8%	731	61,8%
Pulaar	9	2,3%	6	1,5%	10	2,5%	25	2,1%
Seereer	43	11,1%	41	10,3%	47	11,9%	131	11,1%
Autre	0	0,0%	0	0,0%	1	0,3%	1	0,1%
Deux langues ou +	94	24,2%	96	24,1%	105	26,5%	295	24,9%
Ensemble	388	100,0%	399	100,0%	396	100,0%	1183	100,0%

Une seconde série d'informations concernant l'environnement linguistique des élèves porte sur leur degré de familiarité avec le français. Les parents d'élèves ont été interrogés sur le fait de savoir s'ils comprennent le français, s'ils le lisent et s'ils l'écrivent. Leurs réponses à ces trois questions ont été croisées pour définir leur degré d'aisance en français (tableau 10 ci-dessous). On observe que près des trois quarts (74%) des élèves ont des parents qui ne sont pas ou peu à l'aise en français. Ils sont ensuite 17% à avoir des parents moyennement à l'aise. Enfin, seuls 9% ont des parents à l'aise ou très à l'aise en français. Cette tendance générale vaut pour les élèves des trois classes, ceux de CP étant un petit plus nombreux à avoir des parents à l'aise en français.

Tableau 10 : Distribution des élèves selon le degré d'aisance en français de leurs parents

Aisance en français	CI		CP		CEI		Total	
	N	%	N	%	N	%	N	%
Faible	290	74,7%	286	71,7%	297	75,0%	873	73,8%
Moyenne	65	16,8%	68	17,0%	69	17,4%	202	17,1%
Bonne	33	8,5%	45	11,3%	30	7,6%	108	9,1%
Ensemble	388	100,0%	399	100,0%	396	100,0%	1183	100,0%

Il ne suffit pas que les parents soient à l'aise en français pour que les élèves en tirent nécessairement un bénéfice pour leur apprentissage dans cette langue. De plus, les parents ne sont pas les seuls à intervenir sur ce plan dans la vie de leur enfant. En conséquence, il a été demandé aux élèves si quelqu'un à la maison parle français et la fréquence à laquelle cette personne lui parle dans cette langue. Ces informations ont été combinées pour faire ressortir la fréquence des interactions entre les élèves et une personne de leur entourage proche parlant français. Elles sont présentées dans le tableau 11 ci-dessous.

Tableau 11 : Distribution des élèves selon le degré d'interactions avec une personne parlant français à la maison

Interactions en français	CI		CP		CEI		Total	
	N	%	N	%	N	%	N	%
Quelqu'un parle français	136	35,1%	179	44,9%	200	50,5%	515	43,5%
Fréquence des interactions								
Aucune	89	22,9%	97	24,3%	89	22,5%	275	23,2%
De temps en temps	30	7,7%	45	11,3%	59	14,9%	134	11,3%
Régulières	17	4,4%	37	9,3%	52	13,1%	106	9,0%
Personne ne parle français	252	64,9%	220	55,1%	196	49,5%	668	56,5%
Total	388	100,0%	399	100,0%	396	100,0%	1183	100,0%

Un peu plus de la moitié des enfants (56%) n'a personne à la maison avec qui parler français. Parmi les 43% d'élèves qui ont une personne parlant français à la maison, 9% interagissent dans cette langue avec cette personne de façon régulière (« quelques fois par semaine » ou « tous les jours »), 11% de temps en temps et 23% jamais. Les proportions d'élèves qui ont des interactions en français avec des personnes de leur entourage augmentent avec le niveau scolaire : de 12% en CI à 21% en CP et à 28% en CEI. Ces chiffres sont intéressants à considérer, dans la mesure où il serait sans doute bénéfique pour les enfants que des interactions en français soient initiées dès leur plus jeune âge.

Une dernière question porte sur le fait de savoir si les enfants regardent des programmes en français à la télévision (tableau 12 ci-dessous). De façon globale, un peu plus de la moitié des enfants (54%) regarde des émissions en français. De plus, leur proportion augmente avec le niveau de scolarisation (48% en CI, 54% en CP et 61% en CEI).

Tableau 12 : Distribution des élèves selon s'ils regardent des programmes TV en français

TV en français	CI		CP		CEI		Total	
	N	%	N	%	N	%	N	%
Oui	188	48,5%	214	53,6%	242	61,1%	644	54,4%
Non	200	51,5%	185	46,4%	154	38,9%	539	45,6%
Ensemble	388	100,0%	399	100,0%	396	100,0%	1183	100,0%

En résumé, l'échantillon comporte une majorité d'élèves monolingues (75%), qui utilisent surtout le wolof (62% de l'échantillon contre 11% pour le seereer et 2% pour le pulaar). Un quart des élèves utilise au moins deux langues dans leur vie quotidienne. En ce qui concerne la pratique du français, 74% des enfants ont des parents n'ayant qu'un faible niveau de maîtrise de cette langue. De plus, pour 56% des enfants, personne ne parle français à la maison. Toutefois,

cette proportion diminue en fonction du niveau scolaire (de 65% en CI à 55% en CP et à 49% en CEI). À l'inverse, la proportion des enfants qui regarde la télévision en français (54% en moyenne) augmente du CI (un peu plus de 48%) au CP (54%) et au CEI (61%).

2.3. Protocole d'évaluation

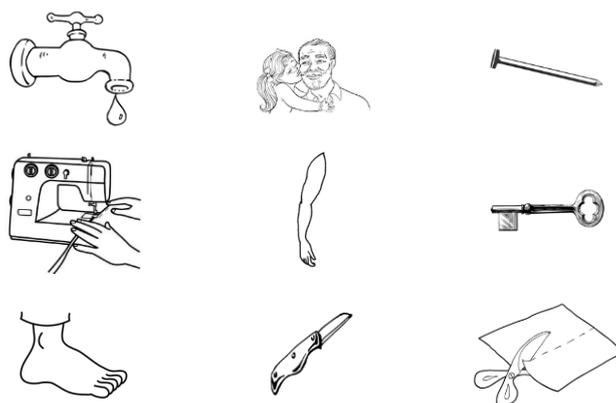
Les tests de vocabulaire et de lecture développés pour la présente étude, qui nécessitent tous une passation individuelle, ont été élaborés en français mais administrés en langue nationale (exceptés les mots dont la connaissance était à évaluer, qui ont été donnés en français). Il est à noter que, pour connaître l'incidence du niveau initial de vocabulaire oral sur les capacités ultérieures de lecture, il aurait été nécessaire d'avoir des informations sur ces mêmes capacités dans la LN des élèves. Les tests auraient donc dû être adaptés aux trois LN du Sénégal ciblées par l'étude, ce que nous n'avons malheureusement pas pu faire pour ceux évaluant le vocabulaire.

2.3.1. Les tests développés

2.3.1.1. Niveau de vocabulaire oral

Le niveau de vocabulaire oral a été évalué par quatre tâches. Le premier test, comme expliqué dans le cadre théorique, est très souvent utilisé avec de jeunes enfants : la désignation de l'image (parmi plusieurs) qui correspond à un mot donné oralement par la personne en charge de l'expérimentation. L'épreuve développée pour la présente étude contient 4 planches de 9 images présentées successivement, l'élève devant trouver l'image correspondant au mot qui lui est donné à l'oral : par exemple, les mots « robinet », puis « bras », puis « coudre » et enfin « clé » pour la planche ci-dessous (graphique 2).

Graphique 2 : Exemple de planche utilisée pour la désignation d'image dans le test



Le choix des mots pour ce test s'est fait à partir de quatre sources de données : celles de l'épreuve de l'EVIP (Vocabulaire en images de Peabody pour les enfants de 2 ans et demi à 6 ans), celles du TVAP (test de Vocabulaire Actif et Passif pour les enfants de 3-5 ans et de 5-8 ans)²⁶, la liste des mots les plus fréquents dans les manuels français de CP (d'après Manulex) et celle des mots inclus dans le curriculum de l'éducation de base du Sénégal (annexe 2 du

²⁶. EVIP (Échelle de vocabulaire en image du Peabody), Dunn, Thériault-Whalen & Dunn (1993) ; TVAP (Test de vocabulaire actif et passif), Deltour & Hupkens (1980).

curriculum). En plus des critères concernant la fréquence des mots, leur sélection s'est effectuée de façon à intégrer quelques verbes (6), en plus des noms communs (16).

Ce test est présenté dans les annexes A1 (catégorie des mots, fréquence d'après MANULEX, présence dans l'EVIP 2-7 et / ou dans le TVAP 3-8) et B1 (consignes pour la présentation du test, sa passation et son codage, total = 22 points). Il est à noter que, afin de ne pas prolonger inutilement cette épreuve pour des enfants n'ayant qu'une connaissance très limitée de la langue française, un critère d'arrêt a été fixé pour ceux qui n'ont donné qu'une réponse correcte sur les 5 items de la première planche d'images. Le niveau de vocabulaire a aussi été évalué par une tâche dans laquelle il était demandé à l'enfant de montrer des parties de son corps (*son nez, sa bouche, sa main droite*). Un point a été attribué à chaque réponse correcte (soit 2 points quand l'enfant montrait correctement sa main droite). Une autre épreuve de même type impliquait la réalisation d'actions simples (*rire, tousser*) ou plus complexes (*tracer une ligne avec son doigt*). Un critère d'arrêt a été fixé après les deux premiers items. Ce niveau a enfin été évalué par une tâche de production : le rappel des jours de la semaine (*Quel est le jour après lundi ? Quel est le jour après mardi ? etc.*).

2.3.1.2. Niveau de prélecture et de lecture

Pour évaluer l'impact du vocabulaire oral sur la lecture, il faut recueillir des données sur le niveau de lecture des enfants. Pour évaluer rapidement ce niveau chez les enfants en fin de CP (ou en début de CE1), nous avons développé pour la présente étude un test de lecture à haute voix en une minute et trente secondes. Dans la mesure où la cible initiale de cette étude est le français, en tant que L2, et trois LN du Sénégal, les items sont des syllabes (liste 1), des mots des LN cible (liste 2) et des mots français fréquemment utilisés dans ces LN (liste 3). Ces items sont courts (2 à 5 lettres, 1 à 2 syllabes) et réguliers au niveau des correspondances graphème-phonème. La plupart des syllabes sont communes aux français et aux trois LN cibles de l'étude. Pour les enfants les plus jeunes (début CI et CP), un test de prélecture de syllabes (voyelle et consonne-voyelle) a été développé. Afin de permettre des comparaisons entre les résultats des enfants des trois niveaux scolaires, quatre items sont communs aux tests de prélecture et de lecture. Ces deux tests sont présentés dans les annexes A3b et B3b.

2.3.2. Les caractéristiques des mots intégrés dans les tests

Les mots sélectionnés appartiennent à différentes catégories grammaticales : des noms, des verbes ainsi que des adverbes indiquant une relation spatiale (devant, derrière, près, loin...), les repères dans l'espace étant essentiels pour la compréhension du langage. La sélection des mots français pour ces deux séries de tests s'est faite de façon à :

- Éviter les effets *plancher* : vu le niveau scolaire des enfants et le fait que le français est, pour la plupart d'entre eux, une L2, les mots français sélectionnés sont parmi les plus fréquents de MANULEX niveau CP²⁷.
- Assurer toutefois une certaine variabilité dans leur *fréquence* (items de différents niveaux de difficulté).

²⁷. Manulex (cf. Lété, Sprenger-Charolles & Colé, 2004) est une base de données qui contient les mots issus de 54 manuels français du primaire, soit 48 900 formes orthographiques (*chantons, chanteraient...* ; *bateau, bateaux...*) ce qui correspond à 23 900 « lemmes » (les entrées de dictionnaires : chanter, bateau...).

- . Éviter les items ambigus, en particulier les mots français qui ont plusieurs sens dans la LN des enfants.
- . Ne pas intégrer trop d'items contenant des *phonèmes* spécifiques au français (entre autres, les *voyelles* nasales, le /y/ de « lu » et le /v/ de « vol »)²⁸.

Des indications plus précises sur les items sélectionnés, ainsi que sur les consignes de passation et de cotation des tests, sont présentées en annexe.

²⁸. Pour les différences entre langues, nous avons consulté, entre autres, le « Word Atlas of Language Structure » (WALS https://wals.info/languoid/lect/wals_code_wlf)

3. RESULTATS

La suite du rapport est consacrée à la description et à l'analyse des résultats obtenus aux tests qui ont été administrés. Une première partie cible les résultats obtenus à chaque item du test de vocabulaire oral en français, puis à ceux du test de prélecture ou de lecture (Partie 3.1). Les informations sur les items étant denses, elles nécessitent d'être synthétisées, pour d'une part, mettre en lumière le score des élèves dans les deux parties du test (Partie 3.2), et d'autre part, analyser ces scores au regard des caractéristiques des élèves et de leur environnement familial, ainsi que pour les élèves scolarisés en CP et en CEI, du contexte et des activités d'enseignement-apprentissage dont ils ont bénéficié l'année précédant l'enquête (Partie 3.3).

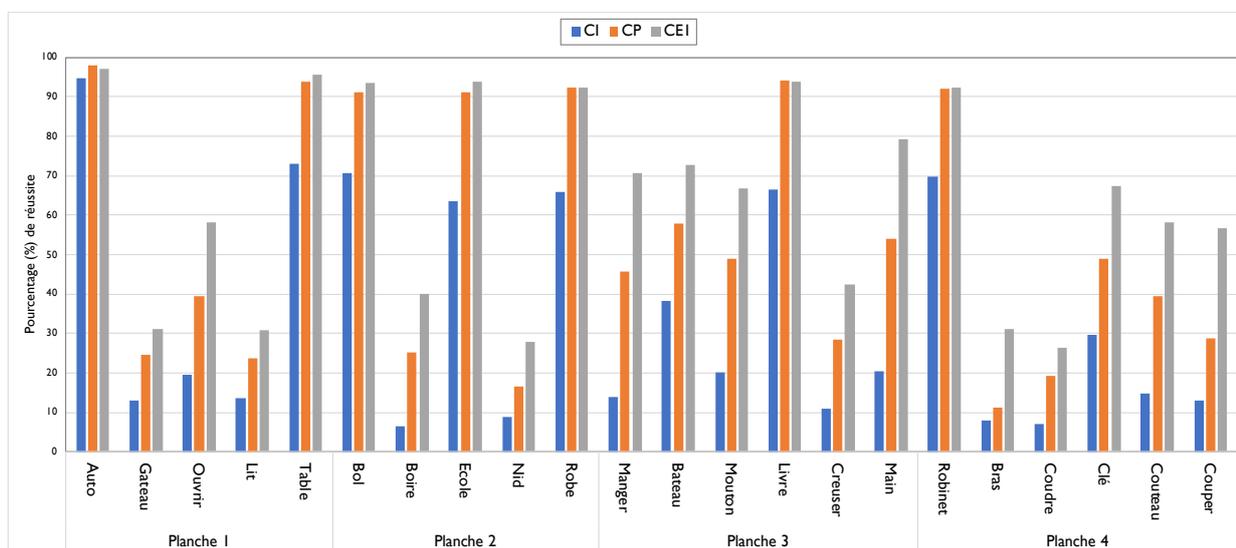
3.1. Les résultats obtenus aux items des épreuves de vocabulaire et de lecture

3.1.1. Épreuves de vocabulaire oral en français

Quatre épreuves ont été développées et administrées aux élèves : une épreuve de désignation d'image, deux épreuves impliquant la réalisation d'actions simples ou complexes et une portant sur la connaissance des jours de la semaine.

L'épreuve de désignation d'image est présentée dans les annexes A1 (catégorie des mots, fréquence, présence dans l'EVIP 2-7 ans et/ou le TVAP 3-8 ans) et B1 (consignes pour la passation et son codage). Pour rappel, afin de ne pas prolonger inutilement cette épreuve pour des enfants n'ayant qu'une connaissance limitée du français, un critère d'arrêt a été fixé. Les résultats sont fournis de façon globale dans le graphique 3 ci-dessous et présentés plus en détails dans les tableaux 13, 14 et 15 ci-après.

Graphique 3 : Pourcentages de réussite des élèves au test de désignation d'images



Un peu moins de 24% des élèves de CI ne vont pas au-delà du critère d'arrêt, contre moins de 4% en CP et CEI. Des pourcentages très bas de réponses correctes ($\leq 10\%$) ne se relèvent qu'en CI sur 2 noms (*nid*, *bras*) et 2 verbes (*coudre*, *boire*). Cette épreuve est donc conservée sans modification pour les analyses subséquentes. Il est aussi à noter que les 5 mots les mieux réussis à tous les niveaux (*auto*, *livre*, *table*, *bol*, *robinet*), ne sont pas les plus fréquents de la base

Manulex, ce qui est surprenant²⁹. Deux autres mots s'ajoutent au CP (*école, robe*) puis trois en CEI (*main, bateau, manger*)³⁰. Ces données signalent la nécessité de disposer de tables de fréquence des mots dans les LN des enfants.

Tableau 13 : Résultats des élèves de CI au test de désignation d'image

Mot	Réussite	Échec	NR	Total	% de réussite	Mot	Réussite	Échec	NR	Total	% de réussite
Planche 1						Planche 3					
Auto	368	16	4	388	94,8	Manger	54	223	20	297	13,9
Gâteau	50	299	39	388	12,9	Bateau	149	138	10	297	38,4
Ouvrir	76	266	46	388	19,6	Mouton	78	207	11	297	20,1
Lit	53	299	36	388	13,7	Livre	258	36	3	297	86,9
Table	28	98	7	388	72,9	Creuser	43	225	29	297	11,1
Échec à 4 items => arrêt pour 91 élèves / 23,5 %						Main	80	201	16	297	20,6
Planche 2						Planche 4					
Bol	274	23	0	297	70,6	Robinet	271	26	0	297	69,8
Boire	25	243	29	297	6,4	Bras	31	242	24	297	8
École	246	48	3	297	63,4	Coudre	28	231	38	297	7,2
Nid	34	230	33	297	8,8	Clé	115	169	13	297	29,6
Robe	256	36	5	297	66	Couteau	57	226	14	297	14,7
(NR : Non Réponse)						Couper	51	222	24	297	13,1

Tableau 14 : Résultats des élèves de CP au test de désignation d'image

Mot	Réussite	Échec	NR	Total	% de réussite	Mot	Réussite	Échec	NR	Total	% de réussite
Planche 1						Planche 3					
Auto	391	7	1	399	98,0	Manger	182	194	8	384	45,6
Gâteau	98	267	34	399	24,6	Bateau	231	143	10	384	57,9
Ouvrir	158	200	41	399	39,6	Mouton	196	177	11	384	49,1
Lit	94	268	37	399	23,6	Livre	375	8	1	384	94,0
Table	374	25	0	399	93,7	Creuser	114	241	29	384	28,6
Échec à 4 items => arrêt pour 15 élèves / 3,8%						Main	215	157	12	384	53,9
Planche 2						Planche 4					
Bol	363	20	1	384	91,0	Robinet	367	17	0	384	92,0
Boire	100	254	30	384	25,1	Bras	45	315	24	384	11,3
École	363	17	4	384	91,0	Coudre	77	275	32	384	19,3
Nid	66	274	44	384	16,5	Clé	195	181	8	384	48,9
Robe	368	16	0	384	92,2	Couteau	158	213	13	384	39,6
						Couper	115	253	16	384	28,8

²⁹. Ordre de fréquence de ces items sur les 22 du test (d'après Manulex, 2004, du plus au moins fréquent) : École (1096); Lit ; Livre (775); Ouvrir; Main (538); Table (499); Gâteau; Couper; Boire; Mouton; Robe (285); Bras; Nid; Auto (172); Clé; Bol (128); Bateau (117); Manger (117); Couteau; Coudre; Robinet (65); Creuser.

³⁰ Une possible explication de la familiarité des enfants avec ces mots français est le fait que ces mots sont utilisés comme emprunts dans les langues wolof, pulaar et seereer.

Tableau 15 : Résultats des élèves de CE1 au test de désignation d'image

Mot	Réussite	Échec	NR	Total	% de réussite	Mot	Réussite	Échec	NR	Total	% de réussite
Planche 1						Planche 3					
Auto	384	11	1	396	97,0	Manger	279	99	4	382	70,5
Gâteau	124	248	24	396	31,3	Bateau	288	93	1	382	72,7
Ouvrir	230	148	18	396	58,1	Mouton	265	112	5	382	66,9
Lit	122	245	29	396	30,8	Livre	372	10	0	382	93,9
Table	379	14	3	396	95,7	Creuser	168	202	12	382	42,4
Échec à 4 items => arrêt pour 14 élèves / 3,5%						Main	314	63	5	382	79,3
Planche 2						Planche 4					
Bol	370	12	0	382	93,4	Robinet	366	16	0	382	92,4
Boire	159	206	17	382	40,2	Bras	123	240	19	382	31,1
École	371	11	0	382	93,7	Coudre	105	259	18	382	26,5
Nid	110	244	28	382	27,8	Clé	267	108	7	382	67,4
Robe	365	15	2	382	92,2	Couteau	231	146	5	382	58,3
						Couper	225	146	11	382	56,8

Dans la seconde série d'épreuves, les élèves devaient réaliser des actions : dans un premier temps, montrer certaines parties de leur corps, puis dans un second temps, réaliser des actions plus complexes. Dans le premier test (tableau 16 ci-dessous), quel que soit le niveau scolaire, les scores les plus élevés sont relevés pour la *main droite* et la *bouche*, suivis par le *nez* et la *langue* et les plus faibles pour le *cou* et le *coude* (moins de 12% de réponses correctes). Ces deux items n'étant pas discriminants, ils ont été éliminés des analyses subséquentes.

Tableau 16 : Résultats des élèves au test de désignation des parties du corps

	Réponse correcte	Réponse partielle	Réponse fausse	Non réponse	Total	% de réussite
CI						
Nez	37	-	230	121	388	9,5
Bouche	112	-	184	92	388	28,9
Langue	32	-	254	102	388	8,2
Cou	5	-	246	137	388	1,3
Main droite	121	62	126	79	388	31,2
Coude gauche	6	6	275	101	388	2,0
Épaule droite	9	5	265	109	388	3,0
CP						
Nez	78	-	194	127	399	19,5
Bouche	186	-	143	70	399	46,6
Langue	60	-	250	89	399	15,0
Cou	15	-	260	124	399	3,8
Main droite	234	94	41	30	399	58,6
Coude gauche	15	2	319	63	399	3,8
Épaule droite	33	4	277	85	399	8,3
CE1						
Nez	141	-	166	89	396	35,6
Bouche	264	-	74	58	396	66,7
Langue	93	-	218	85	396	23,5
Cou	44	-	233	119	396	11,1
Main droite	254	92	39	11	396	64,1
Coude gauche	29	5	329	33	396	7,3

Épaule droite	46	7	287	56	396	11,6
---------------	----	---	-----	----	-----	------

Dans le second test (tableau 17 ci-après), seuls quelques élèves s'avèrent capables d'aller au-delà du critère d'arrêt fixé après les deux premières actions : 17 en CI, 39 en CP et 63 en CEI. Ces résultats suggèrent que les élèves n'ont qu'une compréhension limitée du français oral au-delà du mot. Ce constat est renforcé par le choix de l'équipe qui a jugé préférable de donner à chaque enfant les consignes (sauf le mot test, ou les mots tests) dans sa LN. En conséquence, les résultats de ce second test n'ont pas été pris en compte dans les analyses subséquentes.

Tableau 17 : Résultats des élèves au test de vocabulaire en action

	Réponse correcte	Réponse fausse	Non réponse	Total	% de réussite
CI					
Rire	4	96	288	388	1,0
Tousser	13	87	288	388	3,4
Échec à 2 items => arrêt pour 371 élèves / 95,6. Seuls 17 élèves ont pu poursuivre le test.					
CP					
Rire	17	116	266	399	4,3
Tousser	22	109	268	399	5,5
Échec à 2 items => arrêt pour 360 élèves / 90,2%. Seuls 39 élèves ont pu poursuivre le test.					
CEI					
Rire	37	121	238	396	9,3
Tousser	30	129	237	396	7,6
Échec à 2 items => arrêt pour 333 élèves / 84,1%. Seuls 63 élèves ont pu poursuivre le test.					

Il en est de même pour ceux de la dernière épreuve, dans laquelle les élèves devaient nommer les jours de la semaine. Les résultats (tableau 18 ci-dessous) ne sont pas exploitables dans la suite des analyses pour deux raisons : d'une part, plus de 40% des élèves de CI ne dépassent pas le critère d'arrêt, et d'autre part, les scores de ceux de CEI sont au plafond (91 à 95% de réponses correctes).

Tableau 18 : Résultats des élèves au test de vocabulaire oral : jours de la semaine

	Réponse correcte	Réponse fausse	Non réponse	Total	% de réussite
CI					
Lundi	195	120	73	388	50,3
Mardi	184	141	63	388	47,4
Échec à 2 items => arrêt pour 160 élèves / 41,2 %					
Mercredi	150	69	9	228	38,7
Jeudi	134	77	17	228	34,5
Vendredi	132	75	21	228	34,0
Samedi	172	46	10	228	44,3
CP					
Lundi	347	39	13	399	87,0
Mardi	333	48	18	399	83,5
Échec à 2 items => arrêt pour 41 élèves / 10,3%					
Mercredi	318	38	2	358	79,7
Jeudi	312	39	7	358	78,2
Vendredi	302	42	14	358	75,7
Samedi	341	13	4	358	85,5

	Réponse correcte	Réponse fausse	Non réponse	Total	% de réussite
CEI					
Lundi	377	13	6	396	95,2
Mardi	367	19	10	396	92,7
Échec à 2 items => arrêt pour 17 élèves / 4,3%					
Mercredi	364	13	2	379	91,9
Jeudi	365	11	3	379	92,2
Vendredi	361	13	5	379	91,2
Samedi	370	6	3	379	93,4

3.1.2. Épreuves de prélecture et de lecture

Les résultats des élèves aux épreuves de prélecture ou de lecture sont présentés pour les élèves de CI et de CP dans le tableau 19 ci-dessous, et pour les élèves de CEI dans le tableau 20 ci-après. En CI, en dehors de <o>, <a> et <i>, les scores sont très faibles, surtout quand il faut lire un item consonne-voyelle : seuls 5 à 6% des élèves fournissent des réponses correctes sans les autocorrections, sans changement quand elles sont considérées. En CP, ces pourcentages sont plus élevés (40 à 64% sans les autocorrections, 44 à 65% avec).

Tableau 19 : Résultats des élèves de CI et CP au test de prélecture

	Réponse correcte (RC)	RC après AC*	Réponse fausse	Non réponse	Total	% de réussite sans AC	% de réussite avec AC
CI							
a	164	3	167	54	388	42,3%	43,0%
i	127	6	210	45	388	32,7%	34,3%
é	61	0	262	65	388	15,7%	15,7%
o	183	5	168	32	388	47,2%	48,5%
sa	22	3	280	83	388	5,7%	6,4%
la	22	0	275	91	388	5,7%	5,7%
dé	18	2	267	101	388	4,6%	5,2%
ti	24	0	285	79	388	6,2%	6,2%
CP							
a	359	5	31	4	399	90,0%	91,2%
i	347	7	38	7	399	87,0%	88,7%
é	277	9	101	12	399	69,4%	71,7%
o	381	4	13	1	399	95,5%	96,5%
sa	255	6	100	38	399	63,9%	65,4%
la	206	11	135	47	399	51,6%	54,4%
dé	160	14	189	36	399	40,1%	43,6%
ti	236	7	129	27	399	59,1%	60,9%

(AC : Auto-Correction)

En CEI (tableau 20 ci-après), les pourcentages de réussite sont supérieurs à ceux relevés en CP (62 à 86% sans les autocorrections, 65 à 87% avec). A ce niveau scolaire, les élèves ont aussi eu à lire des mots inventés de deux syllabes ainsi que des mots issus de leur LN ou du français. Pour les mots inventés, les scores varient de 46% à 65% sans les autocorrections (51% à 68% avec), le score le plus faible étant relevé sur le mot le plus long (*sabol*). Concernant les mots des LN, les

scores varient de 52% à 66% sans les autocorrections (56 à 69% avec) et, pour les mots français, ils vont de 51% à 61% sans les autocorrections (53 à 63% avec). En raison des faibles différences entre les scores avec ou sans les autocorrections (de 1,5 à 5% : 5% pour le mot inventé *sabol* et 4% pour *daba*, *bato* et *alima*, qui sont des mots des LN, les mots français étant ceux qui donnent lieu au plus faible pourcentage d'autocorrections : 1,5% pour *kilo*, 2% pour *ami*, *midi*, *moto*, *bal* et 3% pour *Mali*), les analyses statistiques subséquentes n'ont été effectuées que sur les réponses immédiates.

Tableau 20 : Résultats des élèves de CE1 au test de lecture

	Réponse correcte	Réponse correcte après AC	Arrêt à 30 secondes	Réponse fausse	Non réponse	Total	% de réussite sans AC	% de réussite avec AC
sa	339	4	-	39	14	396	85,6%	86,6%
la	335	7	-	37	17	396	84,6%	86,4%
de	306	7	-	64	19	396	77,3%	79,0%
ti	320	8	-	54	14	396	80,8%	82,8%
al	246	13	-	100	37	396	62,1%	65,4%
fo	305	5	-	61	25	396	77,0%	78,3%
tifo	253	10	-	81	52	396	63,9%	66,4%
lato	257	13	-	79	47	396	64,9%	68,2%
tafi	230	13	-	104	49	396	58,1%	61,4%
sabol	182	20	-	132	62	396	46,0%	51,0%
daba	207	16	23	108	42	396	52,3%	56,3%
oto	262	10	30	67	27	396	66,2%	68,7%
bato	216	16	45	82	37	396	54,5%	58,6%
foto	241	11	50	64	30	396	60,9%	63,6%
mati	259	7	55	50	25	396	65,4%	67,2%
alima	213	15	71	65	32	396	53,8%	57,6%
bal	208	9	77	66	36	396	52,5%	54,8%
ami	242	7	78	42	27	396	61,1%	62,9%
midi	216	9	93	45	33	396	54,5%	56,8%
kilo	204	6	101	53	32	396	51,5%	53,0%
moto	211	7	104	43	31	396	53,3%	55,1%
mali	208	12	108	36	32	396	52,5%	55,6%

3.2. Les scores obtenus aux épreuves de vocabulaire et de lecture

Sur la base des résultats obtenus par les élèves aux différentes parties du protocole d'évaluation et conformément aux explications données dans la section précédente, plusieurs scores ont été construits :

- Pour le vocabulaire oral en français :
 - Un score de désignation d'images pour les élèves de chaque niveau ;
 - Un score de désignation des parties du corps (sauf « cou » et « coude ») pour les élèves de chaque niveau ;
 - Un score global de vocabulaire oral en français à partir des items composant les deux scores précédents, pour les élèves de chaque niveau ;

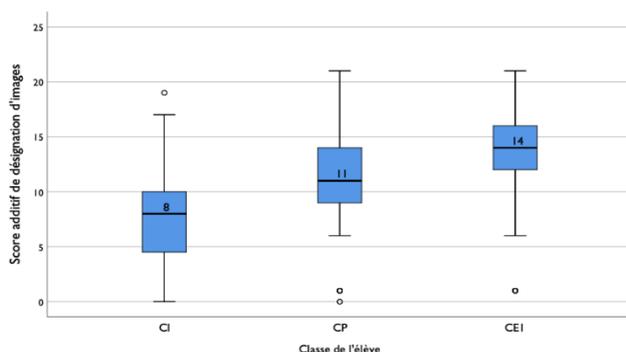
- Pour la prélecture et la lecture (sans tenir compte des auto-corrections) :
 - Un score de prélecture pour les élèves de CI et de CP ;
 - Quatre scores de lecture pour les élèves de CEI :
 - Un score de lecture de 10 mots inventés
 - Un score de lecture de 6 mots en LN
 - Un score de lecture de 6 mots en français
 - Un score global de lecture
 - Un score de prélecture sur les quatre items (« sa », « la », « dé », « ti ») communs aux trois niveaux.

De façon concrète, la construction d'un score résumant les performances des élèves dans chacun des domaines du test, s'est faite selon deux méthodes : (i) la première consiste à sommer de façon directe l'ensemble des scores aux différents items constitutifs d'un domaine et à utiliser ce « score additif » comme performance globale ; (ii) la seconde consiste à conduire une analyse factorielle sur les différents items concernant un même domaine d'acquisitions et à utiliser le premier axe de cette analyse comme variable latente pour caractériser cette performance globale thématique (« score factoriel »). Les analyses de corrélation entre ces deux types de score font apparaître l'existence de similitudes, et ce, pour chacun des domaines traités dans les trois niveaux. Si les scores additifs, plus intuitifs, sont utilisés pour présenter les résultats des élèves, les scores factoriels sont retenus, dans un souci de fiabilité et de comparabilité, pour les autres analyses (ils ont été standardisés avec une moyenne de 100 et un écart-type de 15).

3.2.1. Vocabulaire oral en français

Le niveau de vocabulaire oral en français des élèves est appréhendé à partir des résultats qu'ils ont obtenus aux items des épreuves de désignation d'images et des parties du corps. Le graphique 4 ci-dessous fournit les statistiques descriptives pour les trois classes au test de désignation d'images. L'interprétation de ces graphiques (appelés *boite à moustache*) se fait de bas en haut : pour chaque score, (i) un premier crochet indique la valeur minimale de la distribution, (ii) un trait qui forme le bas de la boîte donne la position du quartile I (25% de la distribution), (iii) un trait au milieu de la boîte fournit la médiane (la valeur qui partage la distribution en deux parties égales), (iv) un trait qui forme le haut de la boîte indique où se situe le quartile 3 (75%) et (v) un dernier crochet indique la valeur maximale de la distribution. Enfin, les ronds identifient des valeurs extrêmes, éloignées de 1,5 à 3 longueurs de boîte. La boîte et les moustaches sont d'autant plus étendues que la dispersion de la distribution est grande.

Graphique 4 : Description des scores de désignation d'images

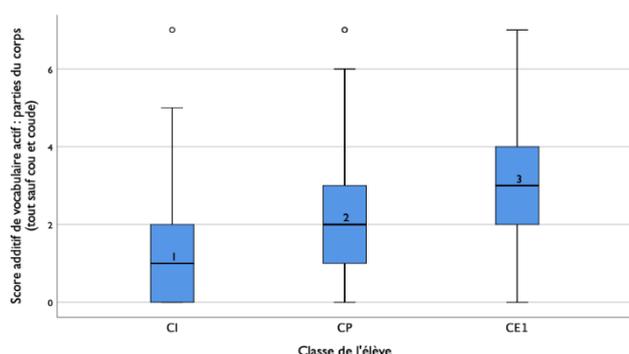


Score de désignation d'images			
	CI	CP	CEI
Moyenne	7,3	11,4	13,6
Écart-type	4,3	3,6	3,9
Min-Max	0-19	0-21	0-21
Percentiles			
25	4,3	9	12
50	8	11	14
75	10	14	16

Le score de désignation d'images est plus élevé à mesure qu'on monte dans les niveaux. En effet, le score moyen des élèves est 7 en CI, 11 en CP et 14 en CEI. En outre, d'une part, les 25% les plus faibles ont obtenu un score compris entre 0 et 4 en CI, entre 0 et 9 en CP et entre 0 et 12 en CEI ; d'autre part, les ronds dans le graphique indiquent qu'en CP et en CEI, seuls quelques élèves ont obtenu des scores vraiment bas. A l'inverse, les 25% des élèves qui ont eu les meilleurs résultats à cette partie du test, ont des scores compris entre 10 et 19 en CI (19 pour seulement deux élèves), entre 14 et 21 en CP et entre 16 et 21 en CEI.

La même tendance se retrouve pour le test de désignation des parties du corps (graphique 5 ci-dessous). Les moyennes vont de 1 en CI, à 2 en CP et à 3 en CEI. La variabilité est moins forte que celle concernant le test de désignation d'images (échelle comprise entre 0 et 7). Le score des 25% des élèves les moins bons est de 0 en CI, compris entre 0 et 1 en CP et entre 0 et 2 en CEI. A l'autre extrême, parmi les 25% qui ont obtenu les meilleurs résultats, le score est compris entre 2 et 7 en CI, entre 3 et 7 en CP et entre 4 et 7 en CEI.

Graphique 5 : Description des scores de désignation des parties du corps



Score de désignation des parties du corps			
	CI	CP	CEI
Moyenne	1,3	2,4	3,0
Écart-type	1,4	1,5	1,6
Min-Max	0-7	0-7	0-7
Percentiles			
25	0	1	2
50	1	2	3
75	2	3	4

Sur la base de ces deux tests (désignation d'images et des parties du corps), nous avons construit deux scores globaux de vocabulaire, un score additif et un score factoriel (selon la procédure présentée en introduction de cette section). Avant d'examiner les résultats des élèves, il est utile d'examiner si les scores que les élèves obtiennent dans une partie du test sont liés à ceux qu'ils obtiennent dans l'autre partie. Par exemple, un élève qui réussit les items de désignation d'images réussit-il de la même façon les items de désignation des parties du corps ? Ceci peut être renseigné à partir d'analyses de corrélation³¹ entre les scores (factoriels) obtenus aux deux parties du test et de façon globale. Le tableau 21 ci-après fournit ces informations. Les chiffres fournis dans les cellules en bleu clair indiquent que les relations entre les scores obtenus dans les deux parties du test sont, à tous les niveaux, plutôt fortes : les élèves ayant obtenu les meilleurs scores dans une partie du test ont plutôt les meilleurs résultats dans l'autre (l'inverse étant également vrai : les élèves qui moins bien réussi une partie du test ont aussi moins bien réussi l'autre). Le fait que les coefficients ne dépassent pas 0,50 signale toutefois que les deux parties du test n'évaluent pas exactement les mêmes compétences. On note également que les coefficients augmentent dans le temps, notamment du

³¹. L'interprétation se fait en fonction de la valeur numérique du coefficient de corrélation (et de son degré de significativité). Rappelons que ce coefficient permet de tester le degré d'association statistique qui existe entre deux variables. Il varie entre -1 et +1 ; l'intensité de cette association est d'autant plus forte que la valeur du coefficient est proche de -1 ou de +1, et d'autant plus faible qu'elle est proche de 0. Lorsque le coefficient est positif, les variables varient dans le même sens et la relation linéaire implicite entre ces variables est croissante. Lorsqu'il est négatif, les variables varient dans le sens contraire, la relation étant décroissante. Notons que l'existence d'une corrélation élevée entre variables n'implique pas l'existence d'une relation causale entre elles.

CP au CEI : les résultats aux deux parties du test tendent donc à être plus homogènes en CEI qu'en CP et qu'en CI. Enfin, les corrélations entre les scores aux deux parties du test et le score global (en bleu foncé dans le tableau ci-dessous) montrent que la partie du test concernant la désignation d'images est celle qui, dans les trois niveaux, apporte la contribution la plus forte à la définition statistique du score global de vocabulaire ; le poids des résultats des élèves à la partie portant sur la désignation des parties du corps augmentant toutefois entre le CI et le CEI.

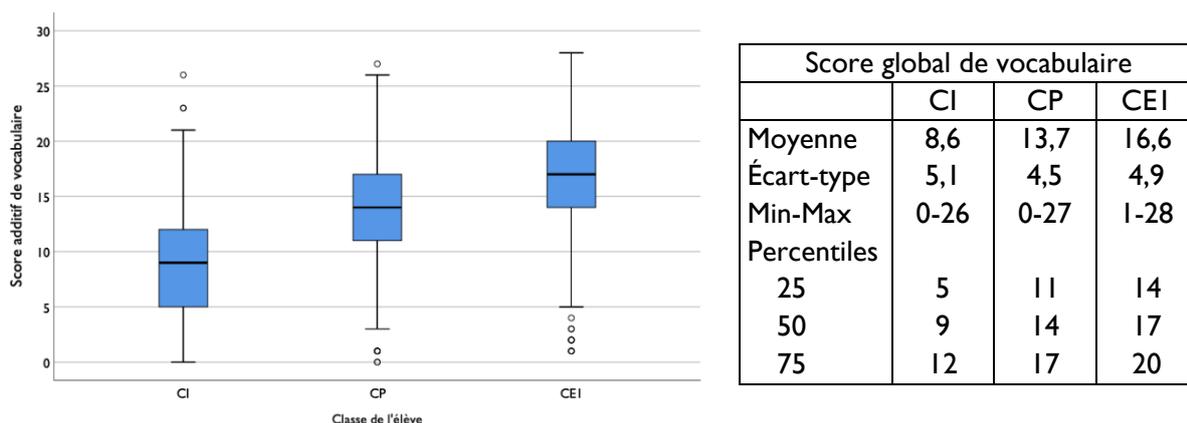
Tableau 21 : Corrélations entre les scores (factoriels) obtenus aux épreuves de vocabulaire

	Désignation d'images	Désignation des parties du corps	Score global vocabulaire
CI			
Score désignation d'images		0,38	0,99
Score désignation parties du corps	-		0,50
Score global vocabulaire	-	-	
CP			
Score désignation d'images		0,39	0,98
Score désignation parties du corps	-		0,56
Score global vocabulaire	-	-	
	Désignation d'images	Désignation des parties du corps	Score global vocabulaire
CEI			
Score désignation d'images		0,46	0,98
Score désignation parties du corps	-		0,61
Score global vocabulaire	-	-	

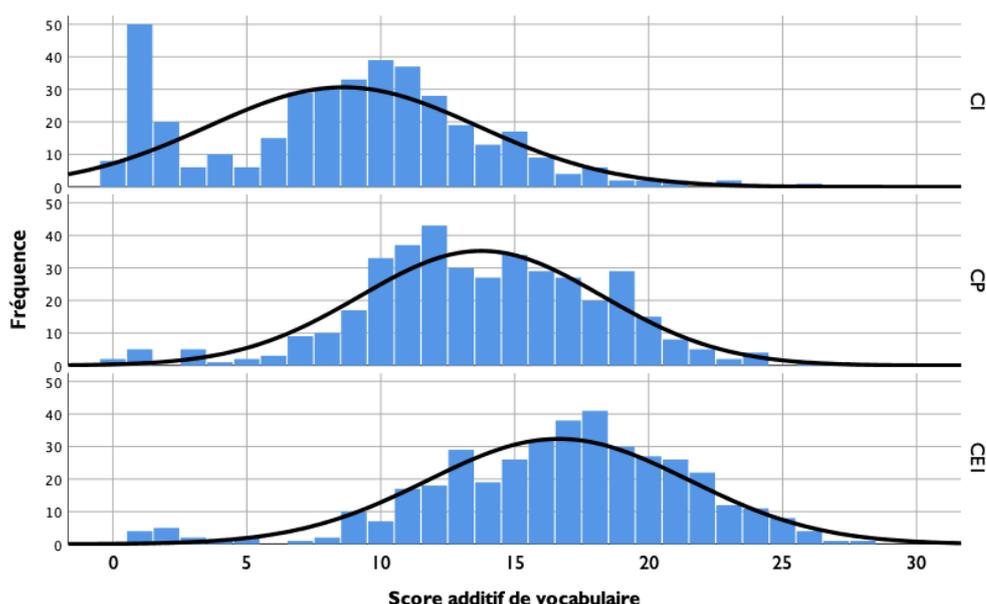
(Tous les coefficients sont significatifs au seuil de 0,001)

Les graphiques 6 et 7 ci-après présentent les statistiques descriptives et la distribution du score global de vocabulaire, déclinées pour chaque niveau. De façon logique, on retrouve les mêmes tendances que dans les analyses précédentes : le score de vocabulaire oral en français est plus élevé à mesure qu'on monte dans les niveaux. En effet, le score moyen s'élève à 9 en CI, à 14 en CP et à 17 en CEI. De plus, les 25% des élèves les moins bons ont obtenu, en CI, un score compris entre 0 et 5 (avec comme le montre le graphique 6, une forte proportion d'élèves ayant 1), en CP, entre 0 et 11 et, en CEI, entre 1 et 14 (en CP et en CEI, seuls quelques élèves ont des scores vraiment bas). A l'inverse, les 25% des meilleurs élèves ont des scores qui varient de 12 à 26 en CI (très peu d'élèves obtenant un score supérieur à 19), 17 à 27 en CP et de 20 à 28 en CEI.

Graphique 6 : Description du score global de vocabulaire



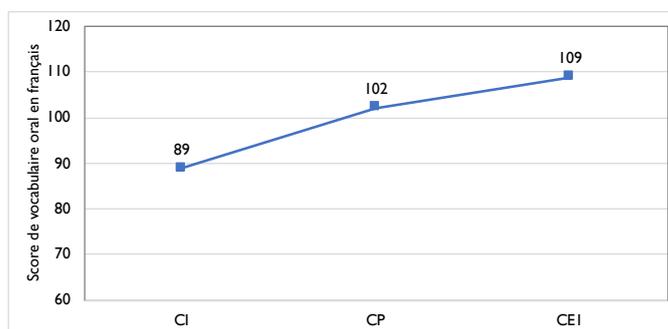
Graphique 7 : Distribution des scores de vocabulaire des élèves des trois niveaux



Jusqu'à présent, nous avons examiné les scores obtenus par les élèves de façon brute et avons mis en lumière le fait que leur niveau de vocabulaire oral en français augmente à mesure qu'on considère des niveaux de scolarité plus élevés. Dans la mesure où les données concernent un échantillon d'élèves et non une population entière, il convient d'aller plus loin dans les analyses et d'examiner si la différence de niveau de vocabulaire oral en français selon la classe dans laquelle les élèves sont scolarisés que nous avons observée est significative (au sens statistique) et si elle l'est, d'identifier son ampleur. Pour cela, nous avons construit un modèle de régression dans lesquels la variable à expliquer est le score obtenu en vocabulaire oral en français (construit sur la base d'une analyse factorielle et standardisé avec une moyenne à 100 et à un écart-type à 15), et les variables explicatives sont la classe dans laquelle les élèves sont actuellement scolarisés (CI, CP ou CEI), ainsi que leurs caractéristiques et celles de leur environnement familial (qui servent ici de variable de contrôle, à raisonner pour des élèves ayant des caractéristiques « égales par ailleurs »). Les estimations fournis par le modèle indiquent qu'il existe bien une différence significative dans le score de vocabulaire oral en français obtenu par les élèves selon la classe dans laquelle ils se trouvent. Pour illustrer son ampleur, nous avons estimé, sur la base des résultats du modèle, les scores moyens obtenus

par les élèves de CI, de CP et de CEI au test de vocabulaire oral en français. Ils sont présentés dans le graphique 8 ci-dessous.

Graphique 8 : Scores moyens estimés de vocabulaire oral en français selon la classe des élèves



On observe une relation croissante entre le niveau de vocabulaire oral en français des élèves et la classe dans laquelle ils sont scolarisés. La différence de niveau s'élève à 13 points entre le CI et le CP et à 7 points entre le CP et le CEI (elle est donc de 20 points entre le CI et le CEI, soit plus d'un écart-type). C'est donc au cours de la première année du cycle primaire que le niveau de vocabulaire oral en français des élèves augmente le plus (sans qu'on puisse à ce stade lier cette augmentation à quel que facteur que ce soit).

3.2.2. Prélecture et lecture

Le niveau de lecture a été évalué par des épreuves différentes selon la classe dans laquelle les élèves sont scolarisés. Ceux de CI et de CP ont passé un test de prélecture, tandis que ceux de CEI ont passé un test de lecture, quatre items (*sa, la, dé et ti*) étant communs aux deux tests. Le tableau 22 ci-dessous fournit la distribution des scores de prélecture des élèves de CI et de CP.

Tableau 22 : Distribution des scores de prélecture des élèves de CI et de CP

	CI		CP	
	N	%	N	%
0	130	33,5%	8	2,0%
1	106	27,3%	7	1,8%
2	64	16,5%	20	5,0%
3	35	9,0%	53	13,3%
4	29	7,5%	51	12,8%
5	5	1,3%	38	9,5%
6	3	0,8%	45	11,3%
7	5	1,3%	65	16,3%
8	11	2,8%	112	28,1%
Total	388	100,0%	399	100,0%
Moyenne	1,6	-	5,6	-
Écart-type	1,9	-	2,2	-

En CI, un tiers des élèves n'a lu aucun des items du test de prélecture, 27 % n'en a lu qu'un et 16 % n'en a lu que deux. Au total, plus des trois quarts (77 %) des élèves ont donc un niveau très limité de prélecture. En CP, ce n'est le cas que pour 9% des élèves, alors que 44% les ont lus tous ou presque tous les items. De façon complémentaire, nous avons cherché à savoir si la

différence observée dans les scores de prélecture des élèves de CI et de CP est significative. Les estimations fournies par un modèle de régression visant à expliquer le score obtenu en prélecture par la classe et par les caractéristiques des élèves et de leur environnement familial, indiquent que c'est le cas, la différence moyenne de score de vocabulaire oral en français entre le CI et le CP s'élevant à 22 points (plus d'un écart-type).

En CEI, le niveau de lecture est appréhendé par les scores obtenus (i) aux dix mots inventés, (ii) aux six mots en LN, (iii) aux six mots en français et (iv) de façon globale. Le tableau 23 ci-après fournit les scores pour ces trois sous-listes. Le pourcentage le plus faible d'élèves qui lisent moins de la moitié des items est relevé pour le sous-liste 1, qui contient des mots inventés : 21 %, contre 35 et 41 % pour les sous-listes 2 et 3, qui contiennent des mots respectivement en LN et en français (qui sont toutefois en moyenne plus longs que ceux de la liste des mots inventés). La différence est minime en ce qui concerne ceux qui lisent tous, ou presque tous, les items de ces trois sous-listes (48, 49 et 47 % pour les sous-listes 1, 2 et 3 respectivement).

Tableau 23 : Distribution des scores de CEI aux différentes parties du test de lecture

10 mots inventés			6 mots en LN			6 mots en français		
	N	%		N	%		N	%
0	30	7,6%	0	95	24,0%	0	126	31,8%
1	15	3,8%	1	24	6,1%	1	21	5,3%
2	14	3,5%	2	20	5,1%	2	14	3,5%
3	13	3,3%	3	16	4,0%	3	23	5,8%
4	13	3,3%	4	47	11,9%	4	27	6,8%
5	29	7,3%	5	66	16,7%	5	47	11,9%
6	21	5,3%	6	128	32,3%	6	138	34,8%
7	33	8,3%						
8	37	9,3%						
9	69	17,4%						
10	122	30,8%						
Total	396	100,0%	Total	396	100,0%	Total	396	100,0%

De façon complémentaire, nous avons cherché à savoir si en moyenne, les différences de scores obtenus par les élèves aux trois sous-listes d'items sont significatives. Les analyses menées sur la base des trois scores (remis sur une échelle allant de 0 à 100 ; tableau 24 ci-dessous) indiquent que c'est le cas et que les mots inventés sont en moyenne nettement mieux lus que les autres, les mots en LN étant ensuite un peu mieux lus que ceux en français.

Tableau 24 : Comparaison des scores moyens de CEI aux différentes parties du test de lecture

	Moyenne	Différence	
		Avec le score aux 10 mots inventés	Avec le score aux 6 mots en LN
Score (sur 100) de lecture aux 10 mots inventés	70,0	-	-
Score (sur 100) de lecture aux 6 mots en LN	58,8	+ 11,2	-
Score (sur 100) de lecture aux 6 mots en français	54,3	+ 15,8	+ 4,6

(Les différences sont significatives au seuil de 0,001)

Toutefois, six des dix mots inventés sont plus courts que les mots en LN et en français (ils n'ont qu'une syllabe, contre deux pour les autres) ce qui pourrait expliquer le fait que les élèves obtiennent un meilleur score à ces mots. Pour vérifier cette supposition, nous avons comparé les moyennes des scores comprenant tous les items de même longueur (tableau 25 ci-après). Il ressort des analyses que, à longueur équivalente, la différence entre les scores aux quatre mots inventés et aux six mots en LN n'est pas significative ; elle l'est en revanche entre ces deux scores et celui sur les mots en français. Ainsi, les élèves lisent en moyenne un peu mieux les mots inventés et les mots en LN que les mots en français (la différence s'élevant respectivement à 4 et 5 points sur une échelle de 100).

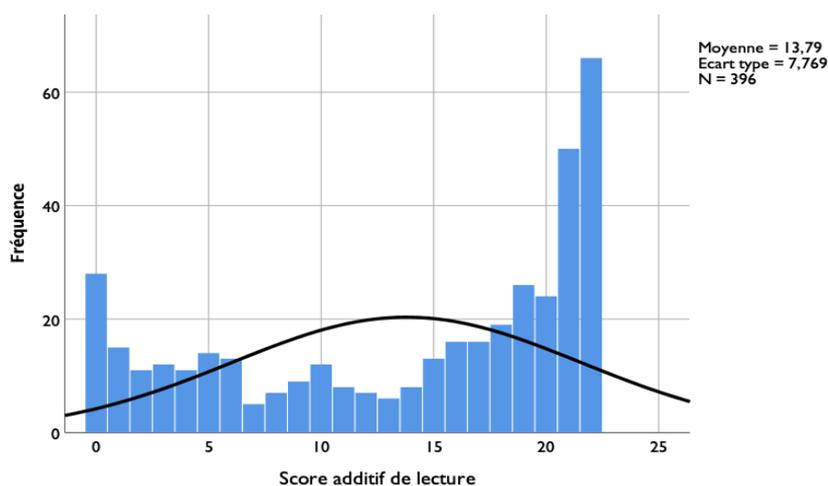
Tableau 25 : Comparaison des scores moyens de CEI aux différentes parties du test de lecture de même longueur (bisyllabiques)

	Moyenne	Différence	
		Avec le score aux 6 mots inventés	Avec le score aux 6 mots en LN
Score (sur 100) de lecture aux 6 mots inventés	58,2	-	-
Score (sur 100) de lecture aux 6 mots en LN	58,8	ns	-
Score (sur 100) de lecture aux 6 mots en français	54,3	3,9 **	4,6 ***

Les seuils de significativité sont les suivants : ns : non significatif ; * = significatif au seuil de 0,05 ; ** = significatif au seuil de 0,01 et *** = significatif au seuil de 0,001. Cette notation est utilisée dans tout le rapport.

En combinant ces trois scores dans un score global, on observe une forte variabilité dans le niveau de lecture des élèves (graphique 9 ci-dessous) : si le score moyen est de 14, les 25% les plus faibles n'ont réussi à lire que 6 mots du test de lecture, alors que les 25% les plus forts, les ont tous, ou quasiment tous, lus.

Graphique 9 : Distribution du score de lecture des élèves de CEI



Enfin, nous avons examiné si les scores obtenus dans une partie du test de lecture sont liés à ceux relevés dans les autres parties. Le tableau 26 ci-après fournit ces informations. Les corrélations entre ces trois sous-tests de lecture passés en CEI sont très élevées, entre 0,76 et 0,84. Les plus élevées sont entre les mots des LN et les mots inventés (0,84) ou ceux qui sont

français (0,83) et les moins élevés sont entre les mots français et ceux qui sont inventés (0,76). Ces résultats suggèrent que les élèves, à ce niveau scolaire, utilisent largement la même procédure (le décodage) pour lire les différents types de mots.

Tableau 26 : Corrélations entre les scores des élèves de CEI aux trois parties du test de lecture

	Score aux 10 mots inventés	Score aux 6 mots en LN	Score aux 6 mots en français	Score global de lecture
Score aux 10 mots inventés	1	0,837	0,761	0,937
Score aux 6 mots en LN	-	-	0,827	0,946
Score aux 6 mots en français	-	-	-	0,917
Score global de lecture	-	-	-	1

(Tous les coefficients sont significatifs au seuil de 0,001)

Jusqu'à présent, les niveaux de prélecture ou de lecture des élèves ont été examinés de façon séparée pour chaque niveau. Les élèves d'un côté de CI et de CP et d'un autre côté de CEI, ayant passé des tests différents, il n'est pas possible de comparer les résultats que nous venons de présenter. On dispose cependant des résultats qu'ils ont obtenus à quatre items communs aux trois niveaux (*sa, la, dé, ti*). La distribution de ce score, déclinée par niveau, est fournie dans le tableau 27 ci-dessous.

Tableau 27 : Distribution du score sur les quatre items communs aux trois niveaux

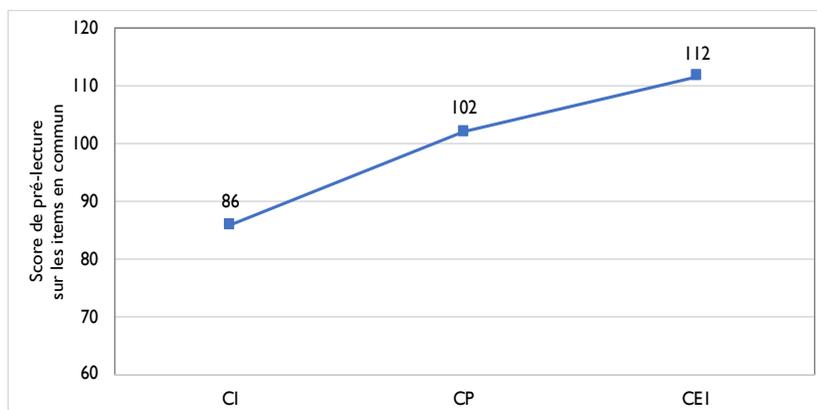
	CI		CP		CEI	
	N	%	N	%	N	%
0	351	90,5%	104	26,1%	33	8,3%
1	17	4,4%	53	13,3%	20	5,1%
2	3	0,8%	44	11,0%	21	5,3%
3	5	1,3%	76	19,0%	50	12,6%
4	12	3,1%	122	30,6%	272	68,7%
Total	388	100,0%	399	100,0%	396	100,0%
Moyenne	0,2	-	2,1	-	3,3	-
Écart-type	0,8	-	1,6	-	1,3	-

Les scores obtenus par les élèves sont très contrastés selon les niveaux. A l'entrée au CI, 90% des élèves ne réussissent pas à lire les quatre syllabes qui leur étaient proposés et 4% ne réussissent en lire qu'une seule ; il n'y a donc que 6% des élèves qui parviennent à lire de deux à quatre des syllabes du test. A l'entrée au CP, si un peu plus d'un quart (26%) des élèves ne lisent aucune syllabe et 13% n'en lisent qu'une seule, ils sont 11% à en lire deux, 19% à en lire trois et 31% à lire toutes les syllabes du test. Enfin, à l'entrée au CEI, un peu plus des deux tiers des élèves (69%) réussissent à lire les quatre syllabes du test et 13% à en lire trois ; 19% des élèves en lisent moins et parmi eux, 8% ne réussissent à lire aucune syllabe du test.

Sur la base de ces résultats, nous avons cherché à savoir si la différence observée entre les niveaux est significative et à en mesurer l'ampleur. Les estimations fournis par un modèle de régression visant à expliquer le score obtenu sur les quatre items communs par la classe (standardisé avec une moyenne de 100 et un écart-type de 15) et par les caractéristiques des

élèves et de leur environnement familial, indiquent que c'est le cas. D'après le graphique 10 ci-dessous, qui fournit une estimation du score moyen obtenu par les élèves de chaque niveau, on observe une différence de 16 points (un peu plus d'un écart-type) entre les élèves de CI et de CP et une différence de 10 points entre ceux de CP et de CEI.

Graphique 10 : Scores estimés de prélecture des items communs selon la classe des élèves



En résumé, le score de vocabulaire oral en français des élèves (standardisé avec une moyenne de 100 et un écart-type de 15) augmente en moyenne en fonction du niveau scolaire. La différence s'élève à 13 points entre le CI et le CP (presque un écart-type) et à 7 points entre le CP et le CEI. Il en est de même en lecture sur les quatre items communs aux différents niveaux scolaires, la différence étant de 16 points (un peu plus d'un écart-type) entre le CI et le CP et de 10 points entre le CP et le CEI. C'est donc au cours de la première année du cycle primaire que le niveau de vocabulaire oral en français et celui de lecture (décodage / identification des mots) augmente le plus. Enfin, la comparaison entre les trois sous-listes du test de lecture passé en CEI signale que les mots inventés ne sont pas mieux lus que les mots en LN, ces deux types d'items étant eux-mêmes significativement mieux lus que les mots en français. Ces résultats indiquent, d'une part, que les élèves de CEI utilisent probablement la même procédure (le décodage) pour lire des mots inventés et des mots fréquents issus de leur LN et, d'autre part, qu'ils lisent mieux les mots fréquents issus de leur LN que ceux issus du français.

3.3. Comment expliquer la variabilité des scores des élèves ?

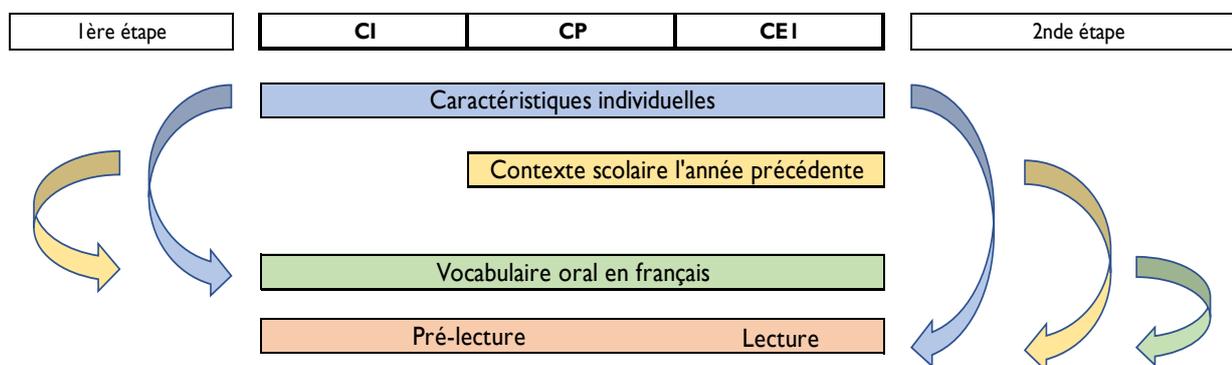
3.3.1. La démarche d'analyse

La présente étude vise une double perspective : (i) une perspective de connaissance du degré de connaissance du vocabulaire oral en français des élèves en début de primaire et dans les premières années du cycle ; (ii) une perspective opérationnelle, pour définir le moment et les modalités d'introduction du français, d'un point de vue tant quantitatif (volume de temps dédié) que qualitatif (contenus des activités), et ainsi appuyer le Ministère de l'Éducation Nationale (MEN) pour finaliser le modèle bilingue de l'enseignement de la lecture qu'il est actuellement en train de définir. Si la première partie de l'étude a permis d'évaluer le niveau de vocabulaire oral

en français et prélecture ou lecture des élèves de CI, CP et CEI, il s'agit, dans cette partie, de répondre au second objectif.

Pour définir le moment et les modalités d'introduction du français, et sachant que l'acquisition du vocabulaire est cruciale pour le développement des compétences langagières et in fine pour l'apprentissage de la lecture, deux pistes d'analyse sont envisagées (graphique 11).

Graphique 11 : Présentation des pistes d'analyse



La première piste vise à expliquer la variabilité des niveaux de vocabulaire oral en français des élèves, en mobilisant les facteurs suivants :

- Les caractéristiques des élèves, ainsi que celles de leur environnement familial : il s'agit d'examiner s'il existe des disparités dans le niveau de vocabulaire oral en français des élèves selon certaines de leurs caractéristiques et à mesurer leur ampleur ;
- Pour les élèves de CP et de CEI, le contexte et les activités d'enseignement-apprentissage dont ils ont bénéficié l'année précédant l'enquête, sachant qu'il peut exister des différences dans :
 - Les modalités d'organisation des classes, sur les plans de la taille de la classe, l'éventuelle mise en œuvre d'un dispositif bilingue particulier, la disponibilité des manuels, la durée de l'année scolaire ou encore les caractéristiques de l'enseignant, que ce soit ses caractéristiques formelles (statut, expérience, formation académique, etc.) ou les perceptions qu'il a de son aisance en français ;
 - Le temps et le contenu des activités dédiés à l'enseignement du vocabulaire oral en français, ce qui permettra d'identifier (i) s'il existe un volume de temps optimal qui devrait être dédié à l'enseignement du vocabulaire oral en français et quel est ce volume, et (ii) s'il existe des types d'activités qui favorisent l'apprentissage du vocabulaire oral en français et quelles sont ces activités.

La seconde étape vise à expliquer la variabilité du niveau de prélecture ou de lecture des élèves, en mobilisant les mêmes facteurs que précédemment :

- Les caractéristiques des élèves, ainsi que celles de leur environnement familial ;
- Le niveau de vocabulaire oral en français des élèves, afin, d'une part, d'analyser si, et dans quelle mesure, ce niveau a un effet sur leur niveau de prélecture et, d'autre part, d'identifier le niveau minimal de vocabulaire qu'il faudrait atteindre pour aborder de façon optimale l'apprentissage de la lecture.
- Pour les élèves de CP et de CEI, le contexte et les activités d'enseignement-apprentissage dont ils ont bénéficié l'année précédant l'enquête, sachant qu'il peut exister des différences dans :

- Les modalités d'organisation des classes, sur les mêmes aspects évoqués plus haut ;
- Le temps et le contenu des activités dédiés à l'enseignement de la prélecture ou de la lecture, pour identifier (i) s'il existe un volume de temps optimal qui devrait être dédié à l'enseignement de la prélecture ou de la lecture et quel est ce volume, et (ii) s'il existe des types d'activités qui favorisent l'apprentissage de la prélecture ou de la lecture et quelles sont ces activités./' ;

3.3.2. Résultats au test de vocabulaire oral en français

3.3.2.1. Par les caractéristiques des élèves et de leur environnement familial ?

Pour chercher à expliquer la variabilité du niveau de vocabulaire oral en français des élèves, une première piste d'analyse consiste à examiner s'il existe des disparités selon certaines des caractéristiques des élèves et de leur environnement, et à mesurer leur ampleur. Nous avons donc construit, pour chaque niveau, un modèle de régression dans lequel la variable à expliquer est le score obtenu en vocabulaire oral en français (construit sur la base d'une analyse factorielle et standardisé avec une moyenne à 100 et à un écart-type à 15), et les variables explicatives les caractéristiques des élèves, ainsi que celles de leur environnement familial. Le tableau 28 ci-après présente les résultats de ces modèles pour les trois niveaux³².

Tableau 28 : Analyses de régression du score de vocabulaire oral en français des élèves selon certaines variables sociales

Variable	Modalités	CI	CP	CEI
Constante		62,76 ***	82,70 ***	90,56 ***
Milieu	Rural	réf.	réf.	réf.
	Urbain	8,84 ***	8,67 ***	7,51 ***
Sexe	Fille	réf.	réf.	réf.
	Garçon	-0,35 (ns)	2,90 *	1,50 (ns)
Age (en année)		2,70 ***	1,05 (ns)	0,59 (ns)
Préscolarisation	Non	-	réf.	-
	Daara		7,66 ***	
Situation socioculturelle de la famille	80% les + défavorisés	réf.	réf.	réf.
	20% les + favorisés	4,91 **	3,42 *	4,47 *
Degré de proactivité pour lire, parler et apprendre des mots en français	Aucune	réf.	réf.	réf.
	Une activité	3,71 *	3,68 **	
	Deux activités	6,27 *	3,85 **	
Regarde la TV en français	Non	réf.	réf.	-
	Oui	6,13 ***	3,75 *	
R ²		26,7%	21,3%	10,1%

Les seuils de significativité sont les suivants : ns : non significatif ; * = significatif au seuil de 0,05 ; ** = significatif au seuil de 0,01 et *** = significatif au seuil de 0,001.

³². Seules les variables significatives ont été maintenues dans les modèles et que des regroupements de plusieurs modalités de variables ont été réalisés lorsque cela s'est avéré pertinent, c'est à dire lorsque plusieurs modalités d'une variable manifestaient, dans une spécification antérieure, un comportement statistique comparable (tant dans leur signe que dans leur ampleur). Ceci s'applique à tous les modèles présentés dans cette étude.

Il ressort tout d'abord, sur la base de la part de variance (R^2) expliquée par les caractéristiques des élèves et de leur environnement familial dans les trois modèles présentés, que le niveau de vocabulaire oral en français des élèves est assez peu dépendant du contexte social et géographique dans lequel ils vivent. Elle s'élève en effet à 27% en CI, 21% en CP et à 10% en CEI. Le fait que les valeurs des R^2 baissent à mesure qu'on considère des niveaux plus élevés suggère que l'emprise des caractéristiques des élèves et de leur environnement familial sur leur niveau de vocabulaire oral en français s'affaiblit à mesure que les élèves progressent dans le système éducatif. Il importe donc de chercher d'autres facteurs que ceux pris en compte ici pour expliquer la variabilité des niveaux de vocabulaire oral en français.

Parmi les caractéristiques des élèves et de leur environnement familial qui concourent à l'explication de la variabilité des performances des élèves au test, la localisation géographique, appréhendée par le milieu de résidence³³, est la caractéristique qui fait le plus de différences sur le niveau de vocabulaire des élèves, et ce, quelle que soit la classe dans laquelle ils sont scolarisés. On observe ainsi que les élèves résidant en milieu urbain obtiennent en moyenne un score supérieur de 8 à 9 points par rapport à ceux habitant en milieu rural.

Concernant les caractéristiques propres aux élèves, on trouve des résultats assez contrastés selon les variables. Ainsi, si le genre des enfants n'a pas d'influence significative sur leur niveau de vocabulaire en CI et en CEI, il fait des différences (de près de 3 points) en CP. A l'inverse, l'âge introduit des différences en CI, le niveau de vocabulaire oral en français augmentant (d'environ 3 points par année) à mesure que les enfants grandissent, mais pas en CP, ni en CEI. La fréquentation du préscolaire améliore également le niveau de vocabulaire oral en français des élèves, mais seulement en CP et lorsque ces derniers ont fréquenté un autre type de préscolaire qu'une daara³⁴. Ils obtiennent alors un score supérieur de près de 5 points par rapport à ceux qui n'ont pas été préscolarisés et à ceux qui ont fréquenté une daara.

Sur le plan de l'environnement familial et en ce qui concerne plus précisément la situation socioculturelle de la famille (qui regroupe sur la base d'une analyse factorielle le niveau d'instruction des parents, leur savoir lire, ainsi que leur profession), on aurait pu penser que le niveau de vocabulaire oral en français des élèves augmenterait de façon linéaire à mesure qu'on considère des situations socioculturelles plus favorisées. Il ressort que ce n'est pas le cas : seuls les élèves venant des familles les plus favorisées sur ce plan ont un meilleur niveau de vocabulaire que tous les autres (score supérieur de 5 points en CI et en CEI et de 3 points en CP).

On trouve un meilleur niveau de vocabulaire oral en français parmi les élèves qui ont des activités telles que parler en français, apprendre des mots en français et lire en français avec quelqu'un de leur entourage (sans que cette personne soit forcément le parent³⁵). Ainsi, le

³³. Dans la mesure où il existe une forte colinéarité entre le milieu de résidence et la zone de résidence (IEF de Gossas, Kaffrine et Kaolack Commune) et compte tenu du fait que le milieu urbain offre en général un environnement plus lettré que le milieu rural (ce qui peut influencer le niveau de vocabulaire des élèves), nous avons décidé de ne conserver que le milieu de résidence dans le modèle.

³⁴ Ceci regroupe la fréquentation d'une structure préscolaire publique, privée ou communautaire. Il n'a pas été possible d'isoler ces différents types de préscolaire du fait du faible nombre d'enfants les ayant fréquentés.

³⁵. Des variables rendant compte de ce type d'interactions entre les parents et leur enfant ont été introduites dans des modèles testés précédemment, sans faire de différence significative. Ces activités sont donc bénéfiques pour l'enfant, quelle que soit la personne qui les mène.

niveau de vocabulaire des élèves s'améliore de 4 à 7 points selon les niveaux et selon le nombre d'activités dont les élèves bénéficient.

On observe une relation positive entre le niveau de vocabulaire oral en français des élèves et le fait qu'ils regardent des programmes en français à la télévision en CI (+ 6 points) ; cette relation s'estompe en CP (+ 3 points) et n'existant plus pour les élèves de CEI.

Notons enfin que certaines dimensions ont été intégrées dans les modèles, sans que leur apport ne fasse de différence significative sur les résultats des élèves. C'est notamment le cas du nombre de langues parlées par les élèves dans leur vie quotidienne³⁶. Ainsi, le fait de vivre dans un environnement monolingue ou bi/multilingue ne contribue pas à rendre compte des différences de niveau de vocabulaire oral en français entre les élèves.

3.3.2.2. Par l'expérience scolaire des élèves l'année précédant l'enquête ?

Pour analyser la variabilité du niveau de vocabulaire oral en français des élèves de CP et de CEI, on dispose, en plus de leurs caractéristiques ainsi que celle de leur environnement familial, d'un certain nombre d'informations sur leur expérience scolaire l'année qui a précédé l'enquête (lorsqu'ils étaient respectivement en CI et en CP). Avant de mener ces analyses, il convient d'abord de décrire les modalités d'organisation des classes dans lesquelles les élèves étaient scolarisés l'année précédant l'enquête (2017-2018), ainsi que le temps et le contenu des activités dédiées au vocabulaire oral en français dont ils ont bénéficié.

Description de l'expérience scolaire des élèves en 2017-2018

On dispose, grâce aux questionnaires qui ont été administrés aux enseignants de ces classes, de nombreuses informations sur les caractéristiques des classes, les caractéristiques des enseignants, les langues utilisées dans le contexte scolaire, ainsi que le temps et le contenu des activités mises en œuvre par l'enseignant. Ces éléments sont examinés de façon successive.

. Les caractéristiques des classes

Une première information concerne la taille des classes dans lesquelles les élèves étaient scolarisés l'année qui a précédé l'enquête. Le tableau 29 ci-dessous fournit le nombre d'élèves par classe, décliné par niveau.

Tableau 29 : Distribution des classes selon leur taille

Taille de la classe	CI		CP		Ensemble	
	N	%	N	%	N	%
< 20	11	28,2%	12	30,0%	23	29,1%
21-30	6	15,4%	9	22,5%	15	19,0%
31-40	7	17,9%	8	20,0%	15	19,0%
41-50	7	17,9%	2	5,0%	9	11,4%
51-60	4	10,3%	5	12,5%	9	11,4%
> 60	4	10,3%	4	10,0%	8	10,1%
Total	39	100,0%	40	100,0%	79	100,0%
Moyenne	39	-	34	-	36	-
Minimum	14	-	9	-	9	-

³⁶. Il n'a pas été possible de distinguer les élèves selon la langue qu'ils parlent (monolingue wolof, pulaar ou seereer) du fait du nombre trop faible d'enfants parlant seulement une de ces trois langues. En conséquence, seul le contexte monolingue versus multilingue a été introduit dans le modèle, mais sans faire de différence significative.

Maximum	124	-	108	-	124	-
---------	-----	---	-----	---	-----	---

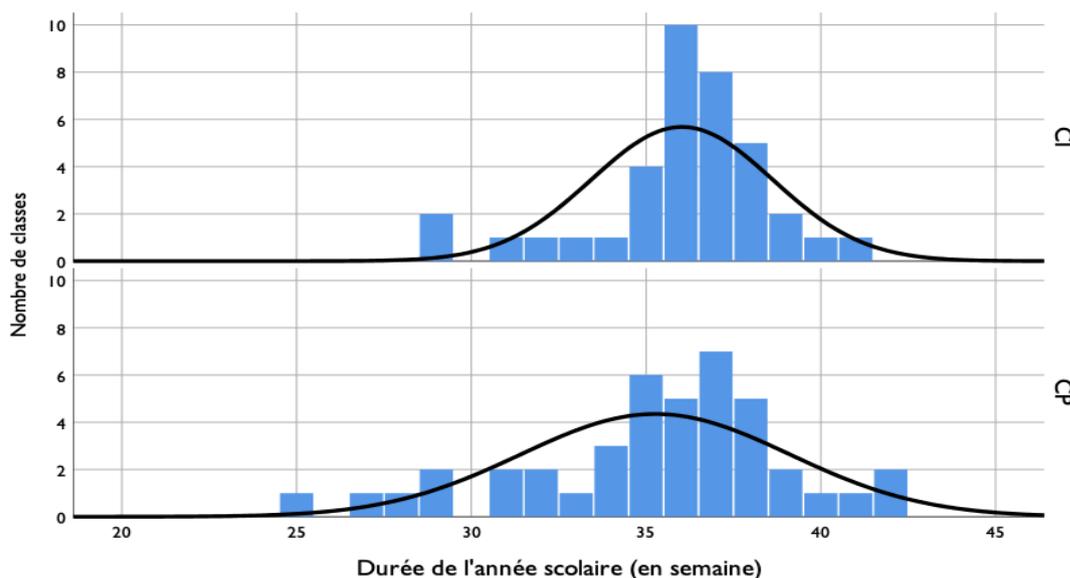
En CI, si le nombre d'élèves par classe s'élève en moyenne à 39, il varie en fait de 14 élèves à plus de 100 et un peu plus d'un quart (28%) des classes compte moins de 10 élèves, 15% en comptent entre 21 et 30, et 20% plus de 50. En CP, le nombre moyen d'élèves par classe est un peu plus faible (36 élèves contre 39) et la variabilité de la taille des classes un peu moins forte : 30% des classes comptent moins de 20 élèves, 19% entre 21 et 30 et 21% plus de 50.

Le temps est une composante essentielle des apprentissages. Il est banal de le dire, mais il faut effectivement du temps pour apprendre. Pour appréhender cette question, les enseignants ont été interrogés sur les dates de début et de fin d'année. Sur cette base, nous avons calculé le nombre de semaines sur lesquelles l'année scolaire 2017-2018 a pris place. Les informations sont données dans le tableau 30 ci-dessous et le graphique 12 ci-après.

Tableau 30 : Début, fin et durée de l'année scolaire

	CI			CP		
	Moyenne	Minimum	Maximum	Moyenne	Minimum	Maximum
Début d'année scolaire	20/10/2017	03/10/2017	03/01/2018	24/10/2017	04/10/2017	02/01/2018
Fin d'année scolaire	01/07/2018	12/06/2018	31/07/2018	02/07/2018	12/06/2018	31/07/2018
Durée de l'année scolaire	36	29	41	35	25	42

Graphique 12 : Distribution des classes selon la durée de l'année scolaire (en semaine)



On trouve des tendances assez similaires dans les classes de CI et celles de CP. Si la rentrée scolaire s'est faite en moyenne aux alentours du 20 octobre, elle s'est en réalité échelonnée entre le début du mois d'octobre et le début du mois de janvier de l'année suivante. L'année scolaire s'est terminée en moyenne début juillet, mais une école l'a terminée mi-juin et une autre fin juillet. Les écarts entre les classes, tant au début qu'en fin d'année, sont considérables. En effet, si parmi les classes de CI, l'année scolaire a duré en moyenne 36 semaines, elle a en réalité varié de 29 à 41 semaines. L'écart est encore plus grand parmi les classes de CP : l'année scolaire s'est étalée en moyenne sur 35 semaines mais elle a en fait duré seulement 25 semaines pour une classe et 42 semaines pour une autre.

Le temps scolaire peut être également varier en cours d'année scolaire. Il a donc été demandé aux enseignants si hormis les vacances scolaires, il y avait eu des interruptions des activités au cours de l'année 2017-2018, et si oui, combien de temps elles avaient duré. Leurs réponses sont dans le tableau 31. Comme précédemment, les tendances sont assez similaires en CI et en CP. Dans environ un tiers des classes, les activités n'ont jamais été interrompues hormis le temps prévu des vacances scolaires. Dans les deux autres tiers, des interruptions ont eu lieu : elles ont duré moins de 15 jours dans 19% des classes, entre 15 jours et un mois dans 23% d'entre elles, entre un et deux mois dans 14% et plus de deux mois dans 11%.

Tableau 31 : Distribution des classes selon la fréquence des interruptions durant l'année

Fréquence des interruptions	CI		CP		Ensemble	
	N	%	N	%	N	%
Jamais	12	30,8%	14	35,0%	26	32,9%
< 15 jours	8	20,5%	7	17,5%	15	19,0%
15 jours à 1 mois	10	25,6%	8	20,0%	18	22,8%
1 à 2 mois	4	10,3%	7	17,5%	11	13,9%
> 2 mois	5	12,8%	4	10,0%	9	11,4%
Total	39	100,0%	40	100,0%	79	100,0%

Les enseignants ont également été interrogés sur le degré de disponibilité du manuel et du livre d'exercices de français (tableau 32 ci-dessous). En CI, si la grande majorité des classes (84%) dispose du manuel de français (sachant toutefois que ce manuel est partagé entre deux élèves dans 19% des classes et entre trois élèves ou plus dans 5% d'entre elles), elles sont « seulement » 41% à disposer du livre d'exercices (qui est partagé entre deux élèves dans 10% des cas). En CP, 80% des classes disposent du manuel de français (partagé entre deux élèves dans 15% des classes et entre trois élèves ou plus dans 5% d'entre elles). Comme précédemment, « seulement » 40% disposent du livre d'exercices. Notons que tous les enseignants (sauf 1 de CI) déclarent utiliser le manuel et le livre d'exercices.

Tableau 32 : Distribution des classes selon la disponibilité du manuel et du livre d'exercices de français

Disponibilité de manuels et livres d'exercices	CI				CP			
	Manuel de français		Livre d'exercices		Manuel de français		Livre d'exercices	
	N	%	N	%	N	%	N	%
Oui, 1 par élève	24	64,9%	12	30,8%	24	60,0%	12	30,0%
Oui, 1 pour 2 élèves	7	18,9%	4	10,3%	6	15,0%	4	10,0%
Oui, 1 pour 3 élèves ou +	2	5,4%	0	0,0%	2	5,0%	0	0,0%
Non	6	16,2%	23	59,0%	8	20,0%	24	60,0%
Total	37	100,0%	39	100,0%	40	100,0%	40	100,0%

En résumé, la taille des classes varie entre moins de 10 élèves à plus de 100 en CI (39 élèves en moyenne) et moins de 20 à plus de 60 en CP (36 en moyenne). La durée de l'année scolaire est également très variable entre les classes : elle va de 29 à 41 semaines, avec une moyenne de 36 en CI et, en CP, de 25 à 42 semaines, avec une moyenne de 35. Il en est de même pour la

durée des interruptions durant l'année, avec également des tendances similaires au CI et au CP : environ un tiers des classes n'ont eu aucune interruption en cours d'année (en dehors des vacances). Parmi les deux tiers restants, l'année scolaire a été interrompue pendant moins 15 jours et entre un mois et plus de deux dans respectivement 19 et 25% des classes. Quant aux supports (manuel et livre d'exercice en français), les manuels sont présents dans 84% des classes de CI et dans 80% de celles de CP, mais partagées entre 2-3 élèves dans 20 à 24% des cas (24% en CI et 20 en CP). Tous les enseignants (sauf 1) disent utiliser les manuels. Seulement 40% des classes disposent d'un livre d'exercice (partagé dans 10% des cas).

Les caractéristiques des enseignants

On dispose d'un certain nombre d'informations sur les enseignants (leur genre, leur langue, leur statut, leur diplôme, leur nombre d'années d'expérience, ainsi que leur niveau de formation initiale et continue). Le tableau 33 ci-après fournit la distribution des enseignants selon leur genre et leur langue maternelle, déclinées selon la classe qu'ils tenaient en 2017-2018.

Tableau 33 : Distribution des enseignants selon certaines caractéristiques personnelles

	CI		CP		Ensemble	
	N	%	N	%	N	%
Nombre d'enseignants	39	-	40	-	79	-
Genre						
Homme	21	53,8%	21	52,5%	42	53,2%
Femme	18	46,2%	19	47,5%	37	46,8%
Langue maternelle						
Pulaar	6	15,4%	3	7,5%	9	11,4%
Seereer	13	33,3%	10	25,0%	23	29,1%
Wolof	18	46,2%	21	52,5%	39	49,4%
Autre	2	5,1%	6	15,0%	8	10,1%

En CI comme en CP, on relève un peu plus d'hommes que de femmes (53% contre 47%). Près de la moitié ont le wolof comme langue maternelle, un peu plus d'un quart (29%) le seereer (33% en CI et 25% en CP), 11% le pulaar (15% en CI et 7% en CP) et 10% une autre langue que les trois qui ont été citées (5% en CI et 14% en CP).

Les données concernant le statut, le diplôme (universitaire et professionnel), le nombre d'années d'expérience et la participation à des actions de formation continue sont présentées dans le tableau 34 ci-après. Près des deux tiers des enseignants sont fonctionnaires (un tiers sont donc contractuels), les fonctionnaires étant plus nombreux en CP qu'en CI (69% contre 57%). Si le nombre d'années d'expérience s'élève en moyenne à 8 et 9 ans, il varie en fait entre 1 et 20 ou 21 années. Plus précisément, le nombre d'année d'expérience est, pour un peu plus d'un quart des enseignants (28%), de moins de cinq ans ; il est de 6 à 10 ans pour plus d'un autre quart d'entre eux (28%), de 11 à 15 ans pour plus d'un tiers (35%), et de 15 ans pour 9%.

Quant aux diplômes, le niveau de 52% des enseignants est au moins équivalent au baccalauréat, 16% ont un niveau supérieur et environ un tiers un niveau correspondant au Brevet de Fin d'Études Moyennes (BFEM) ou au Diplôme de Fin d'Études Moyennes (DFEM). Un peu plus des

deux tiers des enseignants sont titulaires du Certificat d’Aptitude Pédagogique (CAP³⁷) et un quart du Certificat Élémentaire d’Aptitude Pédagogique (CEAP)³⁸. Seuls 2% ont un autre diplôme que le CEAP et le CAP et 6% ne détiennent aucun diplôme professionnel.

Tableau 34 : Distribution des enseignants selon certaines caractéristiques de leur carrière

	CI		CP		Ensemble	
	N	%	N	%	N	%
Nombre d'enseignants	37	-	42	-	79	-
Années d'expérience						
< 5 ans	11	28,2%	11	27,5%	22	27,8%
6-10 ans	10	25,6%	12	30,0%	22	27,8%
11-15 ans	14	35,9%	14	35,0%	28	35,4%
> 15 ans	4	10,3%	3	7,5%	7	8,9%
Moyenne	8,6		8,4		8,5	
Minimum	1		1		1	
Maximum	20		21		21	
Nombre d'enseignants	37	-	42	-	79	-
Statut						
Fonctionnaire	21	56,8%	29	69,0%	50	63,3%
Contractuel	16	43,2%	13	31,0%	29	36,7%
Diplôme académique						
BFEM/DFEM	12	30,8%	13	32,5%	25	31,6%
Baccalauréat	19	48,7%	22	55,0%	41	51,9%
Licence	5	12,8%	4	10,0%	9	11,4%
Maitrise	3	7,7%	1	2,5%	4	5,1%
	CI		CP		Ensemble	
	N	%	N	%	N	%
Nombre d'enseignants	37	-	42	-	79	-
Diplôme professionnel						
Aucun	1	2,6%	4	10,0%	5	6,3%
CEAP	7	17,9%	12	30,0%	19	24,1%
CAP	30	76,9%	23	57,5%	53	67,1%
Autre	1	2,6%	1	2,5%	2	2,5%
Nombre annuel moyen de journées de formation continue (depuis 2010)						
0	1	2,6%	3	7,5%	4	5,1%
1 à 5	12	30,8%	13	32,5%	25	31,6%
6 à 10	14	35,9%	15	37,5%	29	36,7%
11 à 15	8	20,5%	4	10,0%	12	15,2%
> 15	4	10,3%	5	12,5%	9	11,4%
Moyenne	9,2		7,9		8,7	
Minimum	0		0		0	
Maximum	33		32		33	

³⁷. Le CAP est un diplôme réservé aux enseignants. Pour pouvoir le passer, il faut (i) être instituteur décisionnaire, titulaire du baccalauréat et justifier d’au moins un an d’ancienneté, (ii) être instituteur adjoint, ou (iii) être volontaire de l’éducation, titulaire du baccalauréat et justifier d’au moins un an d’ancienneté. Son obtention permet de passer de la catégorie des enseignants adjoints à celle des enseignants titulaires.

³⁸. Le CEAP est un diplôme auquel est associé le grade d’instituteur adjoint (IA). Pour pouvoir le passer, il faut être titulaire du BFEM/DFEM et avoir enseigné pendant au moins un an en tant qu’IA, stagiaire ou volontaire.

Les enseignants ont été interrogés sur le nombre de jours de formation continue dont ils ont bénéficié chaque année, entre 2010 et 2017. Sur la base de leurs réponses, le nombre annuel de journées de formation continue auxquelles ils ont participé ces huit dernières années se révèle être, en moyenne, de 9 jours chaque année (depuis 2010). Toutefois, cette durée est inférieure à 6 jours dans 37% des cas (incluant 5% n'ayant bénéficié d'aucune formation continue). A l'inverse, dans 27% des cas la durée de la formation continue est supérieure à 10 jours (6% d'entre eux ayant bénéficié d'environ un mois de formation par an depuis 2010).

En résumé, les enseignants sont majoritairement des fonctionnaires (57% en CI et 69 % en CP). Ils ont en moyenne entre 8 et 9 ans d'ancienneté (avec toutefois des variations importantes, qui vont de 1 à 20-21 année). La moitié d'entre eux a un niveau d'études équivalent au baccalauréat, 32% ont un niveau inférieur et 16% un niveau supérieur. Presque tous ont un diplôme professionnel. Enfin, entre 2010 et 2017, ils ont bénéficié en moyenne de 9 jours de formation continue par an (plus de 10 jours pour 27% d'entre eux). Seulement 5% n'ont reçu aucune formation continue.

Degré d'aisance en français

De façon complémentaire, il a été demandé aux enseignants dans quelle mesure ils se sentent à l'aise pour lire, écrire, parler, comprendre et enseigner en français. Leurs réponses sont consignées dans le tableau 35 ci-dessous.

Tableau 35 : Distribution des enseignants selon leur perception de leur aisance en français

CI	Lire		Écrire		Parler		Comprendre		Enseigner	
	N	%	N	%	N	%	N	%	N	%
Peu à l'aise	1	2,6%	0	0,0%	0	0,0%	0	0,0%	0	0,0%
Moyennement à l'aise	1	2,6%	2	5,1%	4	10,3%	1	2,6%	1	2,6%
Plutôt à l'aise	8	20,5%	11	28,2%	16	41,0%	13	33,3%	14	35,9%
Très à l'aise	29	74,4%	26	66,7%	19	48,7%	25	64,1%	24	61,5%
Total	39	100,0%	39	100,0%	39	100,0%	39	100,0%	39	100,0%

CP	Lire		Écrire		Parler		Comprendre		Enseigner	
	N	%	N	%	N	%	N	%	N	%
Peu à l'aise	0	0,0%	0	0,0%	0	0,0%	0	0,0%	0	0,0%
Moyennement à l'aise	3	7,5%	4	10,0%	7	17,5%	3	7,5%	6	15,0%
Plutôt à l'aise	10	25,0%	12	30,0%	15	37,5%	12	30,0%	12	30,0%
Très à l'aise	27	67,5%	24	60,0%	15	37,5%	25	62,5%	22	55,0%
Total	40	100,0%	40	100,0%	40	100,0%	40	100,0%	40	100,0%

Quasiment tous les enseignants se déclarent plutôt à l'aise ou très à l'aise pour lire, écrire, parler, comprendre et enseigner en français. Seule une minorité indique l'être peu (un seul enseignant pour la lecture en français) ou moyennement. Parmi ceux qui déclarent être très à l'aise, on note des différences entre les compétences : quel que soit le niveau, les enseignants sont moins à l'aise pour parler en français (49% en CI et 37% en CP) que pour lire (respectivement 74% et 67%), écrire (67% et 60%), comprendre (64% et 62%) ou enseigner (61% et 55%) cette langue. On relève également des différences entre les deux niveaux, les

enseignants de CI étant plus nombreux que ceux de CP à se considérer comme très à l'aise particulièrement pour lire en français (74% en CI et 67% en CP). Les autres domaines de compétences recueillent des chiffres intermédiaires (respectivement pour les deux niveaux, 61% et 55% pour enseigner en français, 67% et 60% pour écrire en français et 64% et 62% pour comprendre le français).

Pour identifier le degré d'aisance en français des enseignants, un score qui combine leurs réponses dans les cinq domaines a été calculé. La distribution de ce score (tableau 36 ci-dessous) varie de 0 pour les enseignants les moins à l'aise à 10 pour les enseignants les plus à l'aise. Si 22% des enseignants se considèrent comme moyennement à l'aise pour lire, écrire, parler, comprendre et/ou enseigner en français (score inférieur ou égal à 5), ils sont en revanche 20% à se percevoir comme plutôt à l'aise dans les cinq domaines (score entre 6 et 7) et 20% se considérer comme très à l'aise (score compris entre 8 et 9). Seulement 38% d'entre eux se perçoivent comme parfaitement à l'aise dans les cinq domaines (score de 10). Ces tendances se retrouvent peu ou prou dans les deux niveaux, les enseignants de CI étant toutefois un peu plus nombreux à se considérer comme parfaitement à l'aise dans tous les domaines (44% contre 32 en CP) et moins nombreux à se considérer moyennement à l'aise (18% contre 25%).

Tableau 36 : Distribution des enseignants selon leur degré d'aisance en français

	CI		CP		Ensemble	
	N	%	N	%	N	%
0-3	1	2,6%	4	10,0%	5	6,4%
4-5	6	15,4%	6	15,0%	12	15,2%
6-7	8	20,5%	8	20,0%	16	20,2%
8-9	7	17,9%	9	22,5%	16	20,3%
10	17	43,6%	13	32,5%	30	38,0%
Total	39	100,0%	40	100,0%	79	100,0%

En résumé, la majorité des enseignants se sent plutôt très à l'aise pour lire, écrire, comprendre et enseigner en français, leur proportion étant toutefois plus faible en CP qu'en CI. Ils sont en revanche moins la moitié (49% en CI et 37 en CP) à se sentir très à l'aise pour parler en français. Dans l'ensemble, peu d'enseignants considèrent que leur niveau de maîtrise du français est plutôt faible : 18% (en CI) et 25% (en CP).

La situation linguistique dans les classes

Les questions suivantes portent sur le contexte linguistique au sein de la classe et plus précisément sur le fait de savoir si les enseignants ont recours à une ou plusieurs autres langues que le français avec leurs élèves, et en premier lieu, dans le contexte d'enseignement-apprentissage (tableau 37 ci-après). Quasiment tous les enseignants déclarent utiliser une autre langue que le français dans la classe (seuls trois enseignants disent ne pas le faire)³⁹. Parmi eux,

³⁹. Parmi les classes de CI, 30 étaient concernées par le programme Lecture Pour Tous en 2017-2018. Dans ces classes, l'enseignant devait introduire le pular, le seereer ou le wolof dans l'enseignement de la lecture. Tous déclarent utiliser dans la classe la langue qui a été choisie pour l'école.

un peu plus des trois quarts disent utiliser une seule langue, le wolof dans la très grande majorité des classes, tandis que 22% déclarent utiliser plusieurs langues nationales.

Tableau 37 : Utilisation d'une autre langue que le français par les enseignants dans la classe

Parle une autre langue dans la classe	CI		CP		Ensemble	
	N	%	N	%	N	%
Oui	38	97,4%	39	97,5%	77	97,5%
Une seule langue	29	76,3%	31	79,5%	60	77,9%
Seereer	3	7,9%	1	2,6%	4	5,2%
Wolof	26	68,4%	30	76,9%	56	72,7%
Plus d'une langue	9	23,7%	8	20,5%	17	22,1%
Non	2	5,3%	1	2,6%	3	3,9%
Total	3	7,9%	1	2,6%	4	5,2%

Les enseignants ont également été interrogés sur le fait de savoir s'ils utilisent une autre langue que le français avec leurs élèves en dehors du contexte d'enseignement (pour annoncer la récréation, reprendre un élève, etc.). Leurs réponses sont consignées dans le tableau 38.

Tableau 38 : Utilisation d'une autre langue que le français par les enseignants hors de la classe

	CI		CP		Ensemble	
	N	%	N	%	N	%
Oui	37	100,0%	37	88,1%	74	93,7%
Une seule langue	28	75,7%	31	83,8%	59	79,7%
Seereer	4	10,8%	1	2,7%	5	6,8%
Wolof	24	64,9%	30	81,1%	54	73,0%
Plus d'une langue	9	24,3%	6	16,2%	15	20,3%
Non	2	5,4%	3	7,1%	5	6,3%
Total	37	100,0%	42	100,0%	79	100,0%

On retrouve les mêmes tendances que celles présentées plus haut concernant les langues utilisées dans la classe. Là encore, quasiment tous les enseignants (sauf cinq) disent utiliser une autre langue que le français en dehors du contexte d'enseignement-apprentissages. Parmi eux, environ 80% déclarent n'utiliser qu'une seule langue (le wolof dans la très grande majorité des cas) ; seuls 20% indiquent qu'ils ont recours à au moins deux langues nationales.

Pour avoir une idée plus précise de la ou des langues utilisées dans et en dehors de la classe, il a été demandé aux enseignants d'estimer la fréquence des interactions en français et la fréquence des interactions en langue nationale. Sur cette base, nous avons croisé leurs réponses pour faire ressortir, de façon un peu brute, trois groupes (tel que cela apparaît dans les zones colorées du tableau 39 ci-après).

- Un premier groupe (en vert), pour lequel les interactions se font de façon plus fréquente en français : 12 classes de CI (31% des classes de CI) et 17 de CP (43%) ;
- Un deuxième groupe (en rose), pour lequel les interactions se font de façon plus fréquente en langue nationale : 16 classes de CI (41%) et 17 de CP (43%) ;
- Un troisième groupe (en jaune), pour lequel les interactions se font de façon quasiment équilibrée en français et en langue nationale : 11 classes de CI (28%) et 6 de CP (15%).

Tableau 39 : Fréquence des interactions en français et en LN dans et en dehors de la classe

CI		Fréquence des interactions dans une ou d'autres langues					
		Très faible	Faible	Moyenne	Forte	Très forte	Total
Fréquence des interactions en français	Très faible				2	3	5
	Faible		3	1	2	1	7
	Moyenne		6	6	2	1	15
	Forte	1	7	2			10
	Très forte	1	1				2
	Total	2	17	9	6	5	39

CP		Fréquence des interactions dans une ou d'autres langues					
		Très faible	Faible	Moyenne	Forte	Très forte	Total
Fréquence des interactions en français	Très faible	1			1		2
	Faible		1		2		3
	Moyenne		8	6	3		17
	Forte		12	2		1	15
	Très forte		3				3
	Total	1	24	8	6	1	40

En résumé, presque tous les enseignants utilisent une autre langue que le français, le plus souvent le wolof, pour leurs échanges avec les élèves, aussi bien dans la classe qu'en dehors (100% en CI et 88% en CP). Environ 20% d'entre eux utilisent une seconde LN.

Les différentes activités d'enseignement apprentissage

Les enseignants ont été interrogés sur la fréquence à laquelle ils ont mis en œuvre différents types d'activités dans leur classe. Le tableau 40 ci-dessous fournit le nombre de minutes consacrés, dans une semaine type, à différentes activités en français pour les élèves qui étaient en CI ou en CP l'année précédente (ceux qui ont été évalués en CP et en CEI). Les enseignants disent avoir consacré, en CI comme en CP, en moyenne, plus de trois heures à enseigner la langue orale et la lecture-écriture (code et compréhension), ainsi qu'à travailler sur des textes et des discours. Ils disent avoir dédié un peu plus de temps à l'enseignement du vocabulaire oral en CP (2 heures) qu'en CI (environ 1,45 heure) et, à l'inverse, moins de temps au travail sur le code en lecture en CP (un peu plus de 1,45) qu'en CI (un peu plus de 2 heures), ce qui est étonnant, tout comme le fait que, quel que soit le niveau, certains disent n'avoir pas consacré de temps à enseigner le vocabulaire, le code (en lecture et écriture) ou la grammaire.

Tableau 40 : Nombre de minutes consacrées chaque semaine à différentes activités

	Moyenne	Écart-type	Minimum	Maximum
CI				
Enseignement de la langue orale	222	143	60	710
Travail autour du texte et du discours	195	153	30	710
Enseignement de la grammaire	87	80	0	300
Vocabulaire oral	107	74	0	300
Enseignement du code pour la lecture	132	84	0	350
Travail de compréhension en lecture	92	48	30	180
Enseignement du code pour l'écriture	98	69	0	350

	Moyenne	Écart-type	Minimum	Maximum
CI				
Travail graphique pour l'écriture	106	68	30	300
CP				
Enseignement de la langue orale	221	124	90	710
Travail autour du texte et du discours	209	124	30	710
Enseignement de la grammaire	96	91	0	300
Vocabulaire oral	120	108	0	600
Enseignement du code pour la lecture	106	73	0	300
Travail de compréhension en lecture	107	71	7	300
Enseignement du code pour l'écriture	93	78	0	300
Travail graphique pour l'écriture	82	59	0	240

Les tableaux 41 à 45 fournissent la fréquence à laquelle les enseignants des élèves vus en CP et en CEI ont mis en place l'année qui a précédé l'enquête (en CI et en CP) différentes activités dans des domaines tels que le vocabulaire, l'enseignement du code en lecture, la compréhension et l'expression orale, etc. En ce qui concerne d'abord le vocabulaire (classement des mots par catégorie : la classe des animaux, des arbres, etc.), le tableau 41 ci-dessous indique que la majorité des enseignants dit pratiquer ce type d'activité quelques fois par semaine, voire tous les jours (en CI, respectivement 69% et 8% ; en CP 57% et 15%). Quelques-uns (3 à 4) disent ne jamais travailler sur le vocabulaire.

Tableau 41 : Distribution des enseignants selon la fréquence des activités de vocabulaire

	CI		CP	
	N	%	N	%
Jamais	3	7,7%	4	10,0%
Une fois par mois	6	15,4%	7	17,5%
Quelques fois par semaine	27	69,2%	23	57,5%
Tous les jours	3	7,7%	6	15,0%
Total	39	100,0 %	40	100,0 %

Le tableau 42 ci-après fournit la fréquence des activités dédiées à la compréhension et à l'expression orale. La majorité des enseignants signale que, quelques fois par semaine, ils racontent des histoires aux élèves, ou leur demandent de raconter un évènement, ou encore de réciter par cœur un poème (en CI, respectivement 59%, 64% et 67% ; en CP, 55%, 70% et 60%). Comme pour le vocabulaire, quelques enseignants disent ne jamais pratiquer ce type d'exercice (2 à 3 ne lisent jamais d'histoire aux élèves et ne leur demandent jamais de raconter un évènement).

Tableau 42 : Distribution des enseignants de CI et de CP selon la fréquence de leurs activités dédiées à la compréhension et à l'expression orale

	L'enseignant lit des récits à haute voix		Il demande à un élève de raconter un évènement		Il demande à un élève de réciter par cœur	
	N	%	N	%	N	%
CI						
Jamais	2	5,1%	3	7,7%	2	5,1%
Une fois par mois	1	2,6%	5	12,8%	8	20,5%
Quelques fois par semaine	23	59%	25	64,1%	26	66,7%
Tous les jours	13	33,3%	6	15,4%	3	7,7%
Total	39	100,0 %	39	100,0 %	39	100,0 %
CP						
Jamais	0	0,0%	2	5,1%	1	2,5%
Une fois par mois	3	7,5%	8	20,5%	12	30,0%
Quelques fois par semaine	22	55,0%	26	66,7%	24	60,0%
Tous les jours	15	37,5%	3	7,7%	3	7,5%
Total	40	100,0 %	40	100,0 %	40	100,0 %

Le tableau 43 ci-après indique la fréquence des activités dédiées à l'apprentissage du code en lecture et en écriture pour les élèves de CI (vus en CP) et de CP (vus en CEI). La majorité des enseignants de CI pratique tous les jours des activités sur le nom, le son et l'écriture des lettres (69%, 72%, 54%) ainsi que sur les phonèmes et les relations graphème-phonème fréquentes (54% et 51%). Celles concernant les relations graphème-phonème dépendant du contexte et la morphologie sont pratiquées plus rarement : quelques fois par semaine par 59% et 44% des enseignants. Les pourcentages relevés en CP sont similaires, avec toutefois moins d'enseignants effectuant tous les jours des activités portant sur les lettres, les phonèmes et les relations graphème-phonème fréquentes (pour le nom, le son et l'écriture des lettres : 60, 55 et 40% ; 40% pour les phonèmes et les relations graphème-phonème les plus fréquentes). Les activités sur les relations graphème-phonème dépendant du contexte sont plus souvent pratiquées qu'au CI (quelques fois par semaine pour 70% des enseignants).

Tableau 43 : Distribution des enseignants selon la fréquence des activités sur le code écrit

	N	%	N	%
CI				
Nombre d'enseignants	39	-	39	-
	Apprentissage du nom des lettres de l'alphabet		Apprentissage du son des lettres de l'alphabet	
Quelques fois / semaine	12	30,8%	11	28,2%
Tous les jours	27	69,2%	28	71,8%
	Apprentissage de l'écriture des lettres de l'alphabet		Sons élémentaires de la langue orale (phonèmes)	
Jamais	0	0,0%	1	2,6%
Quelques fois / semaine	18	46,2%	17	43,6%
Tous les jours	21	53,8%	21	53,8%

	N	%	N	%
CI				
Nombre d'enseignants	39	-	39	-
	Relations graphème-phonème les plus fréquentes		Relations dépendantes du contexte ou de la morphologie	
Jamais	0	0,0%	10	25,6%
Une fois / mois	0	0,0%	3	7,7%
Quelques fois / semaine	19	48,7%	23	59,0%
Tous les jours	20	51,3%	3	7,7%
CP				
Nombre d'enseignants	40	-	40	-
	Apprentissage du nom des lettres de l'alphabet		Apprentissage du son des lettres de l'alphabet	
Jamais	1	2,5%	1	2,5%
Quelques fois / semaine	15	37,5%	17	42,5%
Tous les jours	24	60,0%	22	55,0%
	Apprentissage de l'écriture des lettres de l'alphabet		Sons élémentaires de la langue orale (phonèmes)	
Jamais	1	2,5%	0	0,0%
Une fois / mois	0	0,0%	2	5,0%
Quelques fois / semaine	23	57,5%	22	55,0%
Tous les jours	16	40,0%	16	40,0%
	Relations graphème-phonème les plus fréquentes		Relations dépendantes du contexte ou de la morphologie	
Jamais	0	0,0%	1	2,5%
Une fois / mois	1	2,5%	3	7,5%
Quelques fois / semaine	23	57,5%	28	70,0%
Tous les jours	16	40,0%	8	20,0%

Concernant la grammaire (tableau 44 ci-après), en CI comme en CP, si environ 40% des enseignants disent travailler quelques fois par semaine sur les types de phrases (interrogatives, négatives...), ils sont plus nombreux (environ 45%) à ne jamais aborder cette question. Les exercices de remise en ordre de phrases sont pratiqués plusieurs fois par semaine par 72 à 75% des enseignants de CI et CP respectivement.

Tableau 44 : Distribution des enseignants selon la fréquence des activités de grammaire

	CI		CP	
	N	%	N	%
Nombre d'enseignants	39	100,0 %	40	100,0 %
Travaille sur les types de phrases (interrogatives, négatives...)				
Jamais	18	46,2%	18	45,0%
Une fois par mois	2	5,1%	2	5,0%
Quelques fois par semaine	15	38,5%	16	40,0%
Tous les jours	4	10,3%	4	10,0%
Demande aux élèves de mettre des mots dans l'ordre et d'ajouter celui qui manque				
Jamais	3	7,7%	0	0,0%
Une fois par mois	0	0,0%	2	5,0%
Quelques fois par semaine	28	71,8%	30	75,0%
Tous les jours	8	20,5%	8	20,0%

Le tableau 45 ci-dessous fournit la fréquence des activités dédiées à la conjugaison. Au CI, comme au CP, pratiquement 70% des enseignants proposent des exercices de conjugaison au moins quelques fois par semaine sur le présent (y compris celui des verbes irréguliers) et les moyens d'indiquer le passé et le futur. Les activités autour du passé composé ou du futur des verbes irréguliers sont plus rares. Ce n'est pas trop surprenant pour le futur qui, en français, peut s'exprimer par des expressions non verbales (*je viens demain*). En revanche, c'est regrettable pour le passé composé, qui ne peut s'exprimer que sur le verbe et à partir d'un marquage simple, abordé par la plupart des enseignants : le présent des deux verbes irréguliers les plus fréquents (*être* et *avoir* ; cf. *j'ai vendu ma voiture*, ou *il est parti*). Enfin de très nombreux enseignants disent ne jamais aborder avec les élèves la question des temps verbaux (entre 12 et 61%).

Tableau 45 : Distribution des enseignants selon la fréquence des activités de conjugaison

CI	N		%	
Nombre d'enseignants	39		-	
	Conjugaison des verbes au présent		Verbes irréguliers les + fréquents au présent	
Jamais	8	20,5%	12	30,8%
Une fois / mois	3	7,7%	0	0,0%
Quelques fois / semaine	10	25,6%	14	35,9%
Tous les jours	18	46,2%	13	33,3%
	Moyens d'indiquer le passé		Moyens d'indiquer le futur	
Jamais	6	15,4%	6	15,4%
Une fois / mois	6	15,4%	7	17,9%
Quelques fois / semaine	21	53,8%	21	53,8%
Tous les jours	6	15,4%	5	12,8%
	Verbes irréguliers les + fréquents au passé composé		Verbes irréguliers les + fréquents au futur	
Jamais	24	61,5%	20	61,5%
Une fois / mois	1	2,5%	2	2,6%
Quelques fois / semaine	8	20,5%	12	20,5%
Tous les jours	6	12,5%	5	15,4%
CP	N		%	
Nombre d'enseignants	40		-	
	Conjugaison des verbes au présent		Verbes irréguliers les + fréquents au présent	
Jamais	9	22,5%	9	22,5%
Une fois / mois	0	0,0%	0	0,0%
Quelques fois / semaine	18	45,0%	18	45,0%
Tous les jours	13	32,5%	13	32,5%
	Moyens d'indiquer le passé		Moyens d'indiquer le futur	
Jamais	5	12,5%	5	12,5%
Une fois / mois	4	10,0%	4	10,0%
Quelques fois / semaine	26	65,0%	26	65,0%
Tous les jours	5	12,5%	5	12,5%

CP	N	%	N	%
Nombre d'enseignants	40	-	40	-
	Verbes irréguliers les + fréquents au passé composé		Verbes irréguliers les + fréquents au futur	
Jamais	19	47,5%	19	47,5%
Une fois / mois	3	7,5%	3	7,5%
Quelques fois / semaine	13	32,5%	13	32,5%
Tous les jours	5	12,5%	5	12,5%

En résumé, en ce qui concerne les activités pédagogiques qui sont en lien avec les compétences évaluées dans le cadre du présent projet, les enseignants disent avoir dédié un peu plus de temps à l'enseignement du vocabulaire en CP qu'en CI et, à l'inverse, un peu moins de temps au travail sur le code en lecture en CP qu'en CI, ce qui est étonnant, tout comme le fait que, quel que soit le niveau, certains disent n'avoir pas consacré de temps à enseigner ces deux activités, tout comme celles liées à la compréhension / expression orale. Uniquement certaines des activités liées à l'apprentissage du code sont pratiquées tous les jours par plus de 50% des enseignants : en CI, celles portant sur le nom, le son et l'écriture des lettres (respectivement, 69%, 72%, 54%) ainsi que sur les phonèmes et les relations graphème-phonème fréquentes (54% et 51%) ; en CP celles portant sur le nom et le son des lettres (60 et 55%). Aucune des autres activités n'est travaillé quotidiennement par plus de la moitié des enseignants. Celles concernant le vocabulaire le sont plusieurs fois par semaine par plus de 50% d'entre eux et plus souvent au CI (69%) qu'au CP (57%), tout comme celles portant sur les relations graphème-phonème contextuelles, mais plus souvent au CP (70%) qu'au CI (59%).

Sur la base des réponses fournies par les enseignants, nous avons construit, pour chaque niveau scolaire, des indicateurs standardisés (avec une moyenne de 0 et un écart-type de 1) synthétiques de la fréquence à laquelle les enseignants mettent en place des activités dans les domaines suivants :

- Pôle « Enseignement de la langue orale » à partir de la question portant sur le nombre de minutes consacrée, dans une semaine type, à l'enseignement de la langue orale.
- Pôle « Vocabulaire » à partir des questions portant sur (i) le nombre de minutes consacrée, dans une semaine type, à l'enseignement du vocabulaire oral en français, (ii) la fréquence à laquelle ils travaillent avec les enfants sur le classement des mots par catégorie (animaux...).
- Pôle « Compréhension orale » à partir des questions portant sur (i) le nombre de minutes consacrée, dans une semaine type, au travail autour du texte et du discours en français (lecture et discussions autour d'un récit, etc.), (ii) le nombre de minutes consacrée, dans une semaine type, au travail de compréhension, (iii) la fréquence à laquelle ils lisent des récits à haute voix aux enfants, (iv) la fréquence à laquelle ils demandent aux enfants ce qu'ils ont compris d'un récit qu'ils leur ont lu (qui étaient les personnages, ce qu'ils ont fait, quand, pourquoi...), (v) la fréquence à laquelle ils demandent aux enfants de raconter un événement qui lui est arrivé, (vi) la fréquence à laquelle ils demandent aux enfants de poser des questions sur cet événement et (vii) la fréquence à laquelle ils demandent aux enfants de donner leur avis sur un sujet donné (la nourriture, la ville et la campagne...).
- Pôle « Apprentissage du code pour la lecture » à partir des questions portant sur (i) le nombre de minutes consacrée, dans une semaine type, à l'enseignement du code, (ii) la fréquence à laquelle ils travaillent sur l'apprentissage du son des lettres de l'alphabet, (iii) la

fréquence à laquelle ils travaillent sur les sons élémentaires de la langue orale (phonèmes), (iv) la fréquence à laquelle ils travaillent sur les relations graphème-phonème les plus fréquentes et consistantes, (v) la fréquence à laquelle ils travaillent sur les relations graphème-phonème dépendantes du contexte (par exemple, les deux "s" de « sosie », qui se prononcent différemment selon le contexte) et (v) la fréquence à laquelle ils travaillent sur les relations graphème-phonème pour les marques morphologiques (différences entre écrit et oral ; mots qui se prononcent pareil mais s'écrivent différemment).

- Pôle « Récit par cœur » à partir de la question portant sur la fréquence à laquelle ils demandent aux enfants de réciter par cœur des textes.
- Pôle « Grammaire » : à partir des questions portant sur (i) le nombre de consacrée, dans une semaine type, à l'enseignement de la phrase et de la grammaire en français), (ii) la fréquence à laquelle ils demandent aux enfants de remettre des mots dans l'ordre (par exemple, « Pape Pomme Mange » => « Pape mange une pomme ») et d'ajouter le mot qui manque, (iii) la fréquence à laquelle ils travaillent sur les types de phrases (déclarative, interrogative, négative, impérative).
- Pôle « Conjugaison » à partir des questions portant sur la fréquence à laquelle les enseignants travaillent sur (i) la conjugaison des verbes au présent, (ii) les moyens d'indiquer le passé, (iii) ou le futur, (iv) les verbes irréguliers du français les plus fréquents (être, avoir, faire, dire, aller) au présent, (v) les verbes irréguliers du français les plus fréquents au futur, (vi) les verbes irréguliers du français les plus fréquents au passé composé.

Notons que ces indicateurs sont construits de sorte à ce qu'un score plus élevé corresponde à une mise en œuvre plus fréquente des activités dans la classe.

Résultats des analyses de régression

Après avoir identifié l'existence de différences dans les modalités d'organisation des classes et dans les activités mises en œuvre par les enseignants l'année précédant l'enquête, il importe d'examiner les relations entre la variabilité des conditions et activités d'enseignement-apprentissage et celle des niveaux de vocabulaire oral en français des élèves de CP et de CEI (qui étaient en CI et CP l'année précédant l'enquête). Pour cela, nous avons construit des modèles de régression dans lesquels la variable à expliquer est le score en vocabulaire oral en français (construit sur la base d'une analyse factorielle et standardisé avec une moyenne de 100 et un écart-type de 15) et les variables explicatives sont, en plus des caractéristiques des élèves et de leur environnement familial déjà examinées, les modalités d'organisation des classes et les activités mises en œuvre par les enseignants l'année précédant l'enquête (tableau 46 ci-après).

Une première remarque concerne le pouvoir explicatif des deux modèles (R^2). On observe que la prise en compte, d'une part, des caractéristiques des élèves et de leur environnement familial et, d'autre part, du contexte scolaire et des activités d'enseignement-apprentissage dont ils ont bénéficié l'année précédant l'enquête, n'explique que 31% de la variance des scores des élèves de CP et 23% de la variance de ceux des élèves de CEI. Ces chiffres suggèrent que d'autres facteurs que ceux introduits dans les modèles entrent en compte pour expliquer la variabilité des niveaux de vocabulaire oral en français des élèves.

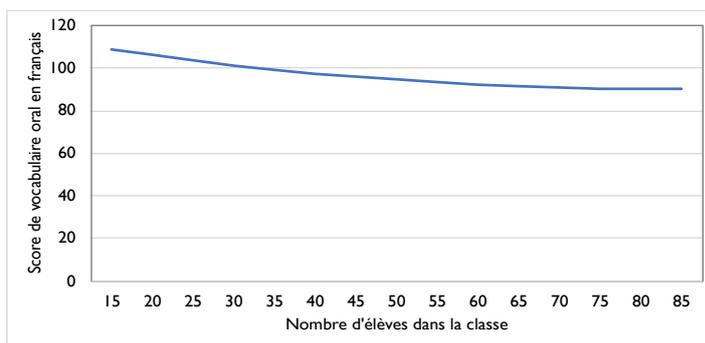
Tableau 46 : Analyses de régression du score de vocabulaire des élèves selon certaines variables sociales et les conditions et activités d'enseignement en 2017-2018 (incluant les caractéristiques des élèves et de leur environnement familial)

Variable		Modalités	CP	CEI
Constante			-322,40 ***	248,20 ***
Caractéristiques de la classe	Taille de la classe		-0,66 ***	-0,14 **
	Taille de la classe ²		0,004 ***	-
	Programme LPT (seulement pour élèves actuellement en CP)	Non	réf.	NC
		Oui	7,31 ***	
	Livres d'exercices de français	Non	-	réf. 5,88 ***
		Oui		
Durée de l'année scolaire (en semaine)		24,48 ***	-8,93 ***	
Durée de l'année scolaire ² (en semaine)		-0,36 ***	0,12 **	
Interruptions durant l'année scolaire	< 30 jours	réf.	-	
	> 30 jours	-7,72 ***		
Caractéristiques de l'enseignant	Niveau de diplôme académique de l'enseignant	BFEM	réf.	
		Bac Licence et +		-8,05 ***
Niveau de diplôme professionnel de l'enseignant	Aucun	-	réf. 5,33 ***	
	CEAP CAP			
Activités d'enseignement-apprentissage	Pôle « Conjugaison »		-2,85 ***	-3,03 ***
	Nombre de minutes pour l'enseignement de la phrase et de la grammaire		-	0,03 ***
	Nombre de minutes pour le travail de compréhension en lecture			-0,11 ***
R ²			31,2%	23,1%

** = significatif au seuil de 0,01 et *** = significatif au seuil de 0,001

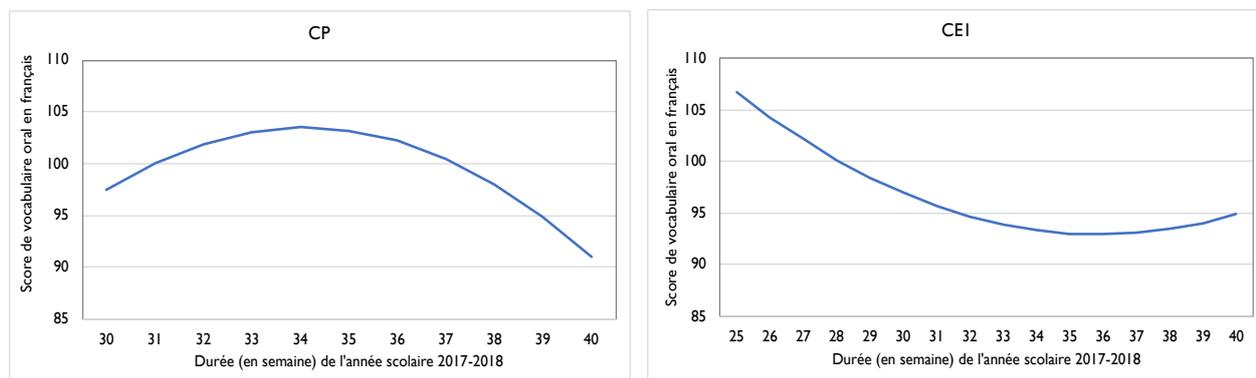
Parmi les caractéristiques des classes qui concourent à l'explication de la variabilité des performances des élèves au test de vocabulaire, on observe tout d'abord une relation négative et significative entre le niveau de vocabulaire oral en français des élèves et la taille de la classe. Rappelons qu'il existe une variabilité assez forte sur ce plan, le nombre d'élèves par classe variant d'un peu moins de 10 à plus de 100 dans les deux niveaux. L'effet de la taille de la classe est certes négatif dans les deux niveaux mais intervient de façon différente : il est linéaire en CEI, un élève supplémentaire dans la classe correspondant à une baisse du score de vocabulaire oral en français des élèves de 0,14 point, alors qu'il ne l'est pas en CP. Pour plus de clarté, nous avons représenté, dans le graphique 13 ci-après, la relation entre les deux variables pour les élèves de CP. On observe ainsi que le niveau de vocabulaire oral en français des élèves diminue lorsque la taille de la classe augmente, avec cependant un seuil qui se situe autour de 50 ou 55 élèves, à partir duquel l'effet négatif s'atténue. Il faut toutefois garder en tête que cet effet est assez modéré dans les deux niveaux.

Graphique 13 : Relation entre le score de vocabulaire oral et la taille de la classe de CI



La durée de l'année scolaire introduit également des différences dans le niveau de vocabulaire oral en français des élèves de CP et CEI. Il existe là encore une grande variabilité dans la durée de l'année scolaire selon les classes : si l'année scolaire 2017-2018 a pris place en moyenne sur 36 semaines pour les élèves de CP et sur 35 semaines sur ceux de CEI, elle a en fait varié de 29 à 41 semaines pour les élèves de CP et de 25 à 42 semaines pour ceux de CEI. Pour une meilleure lisibilité, nous avons représenté, dans le graphique 13 ci-dessous, la relation entre le score de vocabulaire oral en français des élèves et la durée (en semaine) sur laquelle l'année scolaire précédente (CI pour les CP et CP pour les CEI) a pris place.

Graphique 14 : Relation entre le score de vocabulaire oral et la durée de l'année scolaire



La relation entre le score de vocabulaire oral en français des élèves et la durée (en semaine) de l'année scolaire qui a précédé l'évaluation suit des tendances différentes selon les niveaux. Pour les élèves de CP, le graphique fait apparaître qu'une année de CI se déroulant sur une plage allant de 32 à 36 semaines serait optimale pour le niveau de vocabulaire oral en français alors que, pour les élèves de CEI, le chiffre optimal serait, pour une année de CP d'une durée d'environ 25 semaines. Ces chiffres sont plutôt surprenants : on se serait plutôt attendu à ce que l'effet de la durée de l'année scolaire soit similaire dans les deux niveaux. En tout état de cause, le volume de temps dont les élèves ont besoin pour améliorer leur niveau de vocabulaire oral en français semble donc plus important lors de l'année de CI que lors de l'année de CP. Toujours sur le temps scolaire, les analyses montrent que pour les élèves de CP, le fait que les activités scolaires aient été interrompues pendant plus d'un mois pendant leur année de CI a un effet négatif sur leur niveau de vocabulaire oral en français (d'environ 8 points). Cet effet négatif, qui n'existe pas pour les élèves de CEI, ainsi que le résultat portant sur la durée de l'année scolaire suggèrent que les élèves ont besoin, en CI, d'un temps plus long et continu pour améliorer leur niveau de vocabulaire oral en français.

Il faut toutefois garder en tête que ces résultats ne donnent qu'une indication très générale sur le temps scolaire, puisqu'ils concernent seulement la durée sur laquelle l'année scolaire prend place. Il serait intéressant d'analyser plus finalement le temps effectif d'enseignement dont les élèves ont bénéficié, en prenant par exemple en compte le temps de fonctionnement de chaque école, la fréquence et la durée de l'absentéisme des enseignants et des élèves, ou encore le temps d'instruction consacré à chaque activité⁴⁰.

Deux autres caractéristiques des classes influencent le niveau de vocabulaire oral en français des élèves. D'une part, les élèves de CP qui étaient l'année qui a précédé l'enquête dans une classe de CI participant au programme Lecture Pour Tous, et qui ont donc bénéficié d'heures d'enseignement de lecture en langue nationale, ont un niveau de vocabulaire oral en français supérieur (de 7 points) à ceux qui n'ont pas bénéficié de ce dispositif. D'autre part, l'utilisation d'un livre d'exercices en français (qu'il soit partagé ou non) par les élèves de CEI lors de leur année de CP a un effet positif sur leur niveau de vocabulaire oral en français.

S'agissant des caractéristiques des enseignants, de nombreuses variables telles que leur sexe, leur statut, leur nombre d'années d'expérience ou encore le nombre de jours de formation continue dont ils ont bénéficié, ont été introduites dans les modèles, sans que cela fasse de différence significative sur le niveau de vocabulaire oral en français des élèves. En revanche, d'une part, les élèves de CP ayant eu un enseignant titulaire d'une licence ou d'une maîtrise ont un niveau de vocabulaire oral en français inférieur à celui des élèves qui ont eu un enseignant moins diplômé. Ce type de résultat, surprenant au regard de ceux relevés dans les pays développés, est plutôt fréquent dans la littérature sur l'évaluation des acquis des élèves en Afrique⁴¹. D'autre part, les élèves de CEI qui ont eu un enseignant de CP ayant le CAP ont un meilleur niveau de vocabulaire (de l'ordre de 5 points) que les autres (sans formation pédagogique ou titulaires du CEAP).

Concernant enfin les activités d'enseignement-apprentissage dont les élèves ont bénéficié l'année qui a précédé l'enquête, les résultats sont plutôt décevants. En effet, le niveau de vocabulaire des élèves n'est pas influencé par les activités dans ce domaine. Les analyses mettent tout juste en lumière une relation négative, pour les élèves tant de CP que de CEI, entre leur niveau de vocabulaire oral en français et la fréquence des activités dédiées à la conjugaison en classe. Nous reviendrons sur cette question dans la discussion.

Il ressort également des analyses que les élèves de CEI ont un meilleur niveau de vocabulaire oral en français lorsqu'en CP ils ont passé plus de temps à travailler sur l'enseignement de la phrase et de la grammaire, et moins de temps à travailler sur la compréhension pour la lecture. Notons que de façon assez surprenante, si le temps dédié à ces deux domaines a un effet significatif sur le niveau de vocabulaire des élèves, le contenu des activités entrant dans ces deux domaines n'en a pas. Nous reviendrons également sur cette question dans la discussion.

En résumé, le niveau de vocabulaire oral en français diminue quand l'effectif des classes augmente avec un seuil autour de 50-55 à partir duquel cet effet négatif (très modéré aux deux classes examinées, CI pour les élèves vus en CP et CP pour ceux vus en CEI), s'atténue. Le volume de temps dont les élèves ont besoin pour améliorer leur niveau de vocabulaire oral en français est plus important en CI qu'en CP. De plus, le niveau de vocabulaire oral en français est plus élevé (environ 1/2 écart-type) chez les élèves de CP qui étaient en CI dans une classe

⁴⁰. Voir Abdazi (2007).

⁴¹. Mingat & Suchaut (2000); Bernard, Tyab & Vianou (2004) ; Michaelowa & Wechtler (2006).

participant au programme Lecture Pour Tous que chez ceux qui n'ont pas bénéficié de ce dispositif. S'agissant des caractéristiques des enseignants, le niveau de vocabulaire oral en français des élèves de CP (et seulement eux) ayant eu en CI un enseignant titulaire d'une licence ou d'une maîtrise est inférieur à celui des élèves qui ont eu un enseignant moins diplômé. A l'inverse, ce niveau est meilleur chez les élèves de CE1 (et seulement eux) qui ont eu en CP un enseignant ayant le CAP par rapport à ceux qui n'ont pas ce diplôme. Enfin, le temps passé en CP sur des exercices en lien avec la grammaire a un effet positif sur le niveau de vocabulaire oral en français alors que l'effet de celui dédié à la compréhension écrite est négatif. Ces résultats sont d'autant plus surprenants que le contenu des activités dans ces deux domaines n'a pas d'effet.

3.3.4. Résultats aux tests de prélecture et de lecture

Pour définir le moment et les modalités d'introduction du français, et sachant que l'acquisition du vocabulaire est cruciale pour le développement des compétences langagières et, in fine, pour l'apprentissage de la lecture, en plus de l'examen de la variabilité du niveau de vocabulaire oral en français des élèves, il faut examiner les facteurs expliquant la variabilité du niveau de prélecture des élèves de CI et de CP et du niveau de lecture des élèves de CE1.

Une première étape cible les caractéristiques des élèves et de leur environnement familial, de sorte à identifier s'il existe des disparités dans le niveau prélecture ou lecture des élèves selon certaines de leurs caractéristiques et à mesurer leur ampleur. Dans une deuxième étape, on examine si, et dans quelle mesure, le niveau de vocabulaire oral en français des élèves a un effet sur leur niveau de prélecture ou de lecture afin d'identifier s'il existe un niveau minimal de vocabulaire qu'il faut avoir pour aborder de façon optimale l'apprentissage de la lecture. Une troisième et dernière étape s'intéresse aux modalités d'organisation des classes dans lesquelles les élèves de CP et de CE1 étaient scolarisés l'année précédant l'enquête (2017-2018), ainsi qu'au temps et au contenu des activités dédiées à l'apprentissage de la lecture dont ils ont bénéficié. Il s'agit d'examiner les relations entre la variabilité des conditions et activités d'enseignement et celle des niveaux de prélecture ou lecture des élèves.

3.3.4.1. Par les caractéristiques des élèves et de leur environnement familial ?

Pour chercher à expliquer la variabilité du niveau de prélecture des élèves de CI et de CP, nous avons construit, pour chaque niveau, des modèles de régression dans lesquels la variable à expliquer est le score de prélecture obtenu par les élèves (construit sur la base d'une analyse factorielle et standardisé avec une moyenne à 100 et à un écart-type à 15), et les variables explicatives introduites les caractéristiques des élèves et de leur environnement familial. Les estimations de ces modèles sont fournies dans le tableau 47 ci-après.

Les caractéristiques des élèves et de leur environnement familial n'ont que très peu d'emprise sur le niveau de prélecture des élèves. La part de variance (R^2) expliquée par ces variables s'élève à seulement 6% en CI et 13% en CP. Parmi celles qui introduisent des différences, certaines sont communes aux deux niveaux. C'est par exemple le cas de la situation socioculturelle de la famille, les élèves venant des familles moyennes et favorisées sur ce plan ayant un meilleur niveau de lecture que les autres (d'environ 6 points en CI et 3 points en CP). Il en va de même des pratiques de lecture des élèves à la maison, qui ont un effet positif sur le

niveau de prélecture des élèves (de 4 à 6 points selon le niveau et selon que les élèves lisent seuls et/ou avec un adulte). Ce résultat est intéressant car il montre qu'en créant les conditions pour que les enfants regardent ou lisent des livres seuls et/ou avec un adulte, il est possible de les aider à améliorer leur niveau de prélecture.

Tableau 47 : Analyses de régression du score de prélecture des élèves de CI et de CP selon certaines variables sociales et selon le niveau de vocabulaire oral en français

Variable	Modalités	CI	CP
Constante		88,90 ***	82,17 ***
Milieu	Rural	réf.	réf.
	Urbain	-5,34 **	0,54 (ns)
Sexe	Fille	réf.	réf.
	Garçon	-0,10 (ns)	2,77 (ns)
Age (en année)		0,86 (ns)	1,26 (ns)
Fréquentation du préscolaire	0 à 2 ans	-	réf.
	3 ans	-	5,72 *
Situation socioculturelle de la famille	40% les + défavorisés	réf.	réf.
	20% moyens	6,18 ***	3,44 *
	40% les + favorisés		
Lit, parle et apprend des mots en français à la maison	0 à 2 activités	-	réf.
	3 activités	-	5,19 *
Pratiques de lecture à la maison	Aucune	réf.	réf.
	Lit seul ou avec un adulte	6,10 *	4,61 *
	Lit seul et avec un adulte		6,42 ***
R ²		6,0%	12,7%

D'autres caractéristiques ont un effet sur le niveau de prélecture des élèves mais seulement dans un niveau. En CI, le niveau de prélecture s'avère un peu meilleur parmi les élèves résidant en milieu rural (d'environ 5 points par rapport à leurs homologues résidant en milieu urbain, l'écart atteignant même près de 9 points lorsqu'on raisonne pour des élèves ayant un niveau de vocabulaire oral en français équivalent), Cela est surprenant, d'autant plus que cet effet n'est pas significatif en CP. Parmi les élèves de ce niveau, il ressort que ceux ayant été préscolarisés pendant trois ans ont un niveau de prélecture un peu supérieur (près de 6 points) aux autres, cet effet s'estompant cependant lorsqu'on prend en compte le niveau de vocabulaire oral en français. Le fait que les élèves parlent, apprennent des mots et lisent avec quelqu'un de leur entourage a également un effet bénéfique sur leur niveau de prélecture (d'environ 5 points).

Le même type d'analyses a été mené pour les élèves de CEI sur la base de leur score en lecture. Les estimations sont fournies dans le tableau 48 ci-après. Les résultats concernant les élèves de CEI sont assez similaires à ceux concernant les élèves de CP. En effet, les caractéristiques des élèves et de leur environnement familial n'expliquent que très peu (6%) la variabilité du niveau de lecture des élèves. Seule la fréquentation d'une structure préscolaire et le fait de lire à la maison améliorent le niveau de lecture des élèves.

Tableau 48 : Analyses de régression du score de lecture en CEI selon les variables sociales

Variable	Modalités	
Constante		96,46 ***
Milieu	Rural	réf.
	Urbain	2,67 (ns)
Sexe	Fille	réf.
	Garçon	-2,52 (ns)
Age (en année)		0,10 (ns)
Préscolarisation	Non	réf.
	Oui	4,82 *
Pratiques de lecture à la maison	Aucune	réf.
	Lit seul ou avec un adulte	
	Lit seul et avec un adulte	5,38 ***
R ²		6,4%

3.3.4.2. Par le niveau de vocabulaire oral en français des élèves ?

Après avoir vu que les niveaux de prélecture et de lecture des élèves étaient très peu dépendants de leurs caractéristiques et de celles de leur famille, une deuxième piste d'analyse visant à expliquer la variabilité des scores qu'ils ont obtenu cible le niveau de vocabulaire oral en français des élèves. Il s'agit en effet d'analyser si, et dans quelle mesure, ce niveau a un effet sur leur niveau de prélecture ou de lecture, et d'identifier s'il existe un niveau minimal de vocabulaire qu'il est nécessaire que les élèves aient pour aborder de façon optimale l'apprentissage de la lecture.

Pour cela, nous avons construit, pour chaque niveau, des modèles de régression dans lesquels la variable à expliquer est le score de prélecture ou de lecture obtenu par les élèves (construit sur la base d'une analyse factorielle et standardisé avec une moyenne à 100 et à un écart-type à 15), et les variables explicatives introduites : (i) les caractéristiques des élèves et de leur environnement familial ; (ii) le score obtenu par les élèves au test de vocabulaire oral en français, ainsi que ce même score au carré, ce dernier terme visant à tenir compte du fait que la relation entre le niveau de prélecture et le niveau de vocabulaire oral en français peut ne pas être linéaire⁴². Les estimations de ces modèles pour les élèves de CI et de CP sont fournies dans le tableau 49 ci-après.

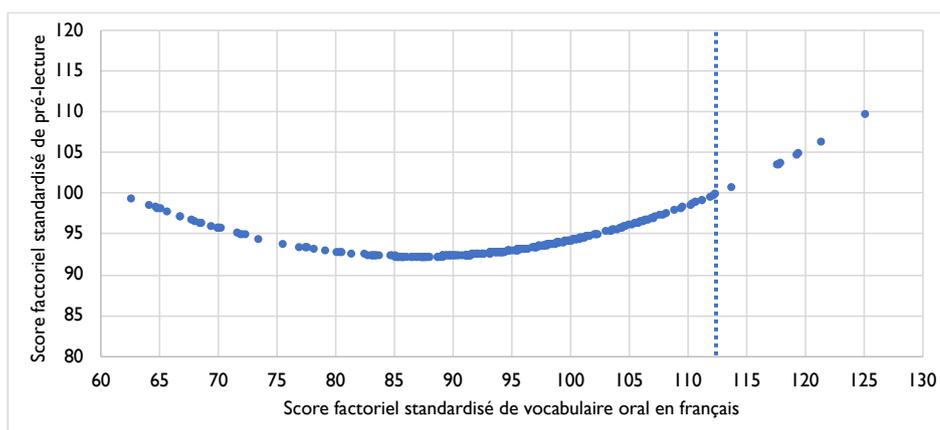
⁴². Il est possible que la relation entre la variable à expliquer (ici, le score de pré-lecture) et la variable explicative (ici, le score de vocabulaire) ne soit pas linéaire. Une indication en est que la représentation graphique de la relation entre les deux variables n'est pas une droite mais une courbe. Dans la mesure où la méthode des moindres carrés ordinaires (utilisée ici) est une méthode de régression linéaire, qui présume que les relations modélisées sont linéaires, il est alors nécessaire de transformer la variable explicative. Plusieurs transformations non linéaires sont possibles (logarithmique, exponentielle, etc.), la plus courante (et celle qui est utilisée ici) est la transformation quadratique, soit l'introduction d'un terme élevé au carré.

Tableau 49 : Analyses de régression du score de prélecture des élèves (CI et CP) selon le niveau de vocabulaire oral en français (avec prise en compte des caractéristiques des élèves et de leur environnement familial)

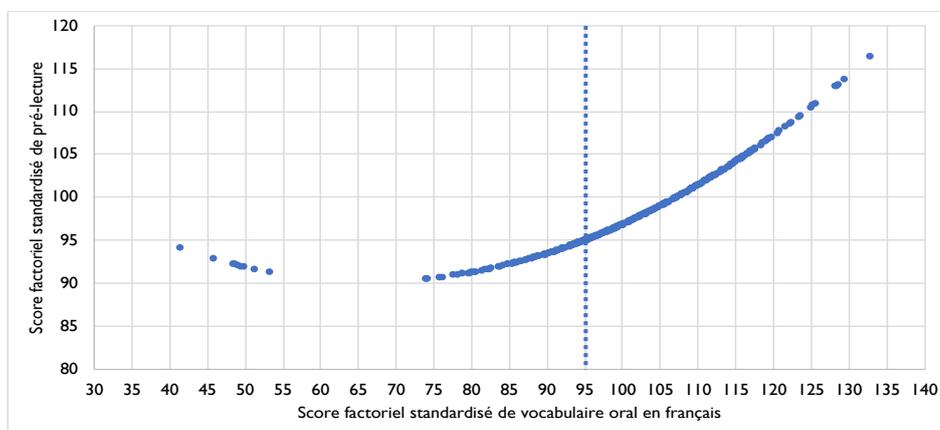
Variable	CI	CP
Constante	177,26 ***	105,27 ***
Score factoriel de vocabulaire	-2,27 ***	-0,80 ***
Score factoriel de vocabulaire ²	0,01 ***	0,01 ***
R ²	16,1%	20,3%

On remarque, en premier lieu, que la prise en compte du niveau de vocabulaire oral en français améliore le pouvoir explicatif des modèles (le R² s'élevait à 6% en CI et à 8% en CEI lorsque seules les caractéristiques des élèves et de leur environnement familial étaient prises en compte). En second lieu, il existe une relation positive et significative entre le niveau de vocabulaire oral en français des élèves et leur niveau de prélecture. Pour plus de clarté, nous avons simulé le score de prélecture obtenu par les élèves selon leur score de vocabulaire oral en français dans les graphiques 15 (pour les élèves de CI) et 16 ci-après (pour les élèves de CP) sur la base des scores effectivement obtenus au test de vocabulaire oral en français par les élèves de l'échantillon.

Graphique 15 : Relation entre les niveaux de vocabulaire oral et de prélecture en CI



Graphique 16 : Relation entre les niveaux de vocabulaire oral et de prélecture en CP



La forme des courbes qui représentent le niveau de prélecture des élèves est similaire dans les deux niveaux : elle stagne au plus bas sous un certain score de vocabulaire oral en français, puis croît de façon nette lorsque les élèves dépassent ce seuil. Il existe donc un niveau minimal de vocabulaire oral en français à avoir pour que cela ait un effet positif sur le niveau de prélecture. Ce seuil varie cependant selon les niveaux : il se situe autour de 110 (à environ un écart-type) à l'entrée en CI et autour de 95 à l'entrée en CP (en deçà de ces scores, le niveau de vocabulaire oral en français ne correspond pas à une amélioration du niveau de prélecture des élèves). En mettant en regard ces seuils avec les scores obtenus par les élèves de l'échantillon, tels que représentés par les points dans les graphiques, on trouve (i) qu'à l'entrée en CI, seuls 4% des élèves ont un niveau de vocabulaire oral en français supérieur à 110 et (ii) qu'à l'entrée en CP, 59% des élèves atteignent un score équivalent ou supérieur à 95.

Ces résultats nous indiquent deux choses. D'une part, il est nécessaire de chercher à améliorer le niveau de vocabulaire oral en français des élèves en CI, avant de débiter l'apprentissage de la lecture, dans la mesure où le niveau de vocabulaire oral en français a un effet positif et significatif sur le niveau de prélecture. D'autre part, le niveau de vocabulaire oral en français de quasiment tous les élèves à l'entrée au CI se situe en dessous du seuil à partir duquel il a un effet sur leur niveau de prélecture ; le niveau de vocabulaire oral en français est encore dessous du seuil au-delà duquel il existe un effet positif pour le niveau de prélecture pour 41% des élèves à l'entrée en CP. De façon mécanique, 59% des élèves de CP disposent d'un niveau de vocabulaire oral en français qui leur permet de « maximiser » leur niveau de prélecture.

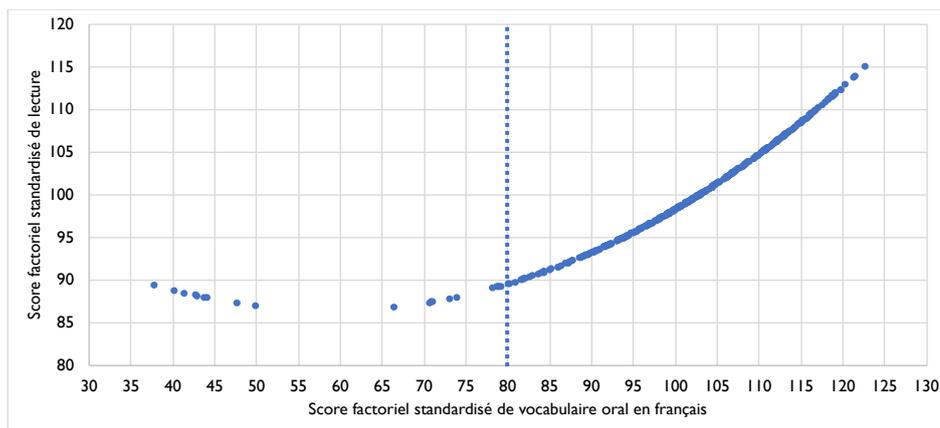
Le même type d'analyses a été mené pour les élèves de CEI, sur la base du score qu'ils ont obtenu au test de lecture, qui est bien plus exigeant. Les estimations sont fournies dans le tableau 50 ci-dessous.

Tableau 50 : Analyses de régression du score de lecture des élèves de CEI selon le niveau de vocabulaire oral en français (avec prise en compte des caractéristiques des élèves et de leur environnement familial).

Variable	Score global de lecture
Constante	109,96 ***
Score factoriel de vocabulaire	-0,83 ***
Score factoriel de vocabulaire ²	0,01 ***
R ²	21,7%

Une nouvelle fois, la prise en compte du niveau de vocabulaire oral en français améliore le pouvoir explicatif des modèles (de 6% lorsque seules les caractéristiques des élèves et de leur environnement familial étaient prises en compte à 22% lorsque le niveau de vocabulaire oral en français des élèves est intégré dans le modèle). De plus, il existe une relation positive et significative entre le niveau de vocabulaire oral en français des élèves et leur niveau de prélecture. Le graphique 17 ci-dessous représente la relation entre les niveaux de lecture et de vocabulaire oral en français des élèves de CEI. Le niveau de lecture des élèves augmente fortement une fois qu'ils ont atteint un certain niveau de vocabulaire oral en français, qui se situe autour de 80. En mettant en regard ce seuil avec les scores obtenus par les élèves de l'échantillon, tels que représentés par les points dans les graphiques, on trouve que seulement 7% des élèves ont un score de vocabulaire oral en français inférieur à ce seuil. La grande majorité d'entre eux (93%) dispose donc d'un niveau de vocabulaire oral en français suffisant pour que cela se traduise de façon positive sur leur niveau de lecture.

Graphique 17 : Relation entre les niveaux de vocabulaire oral et de lecture en CEI

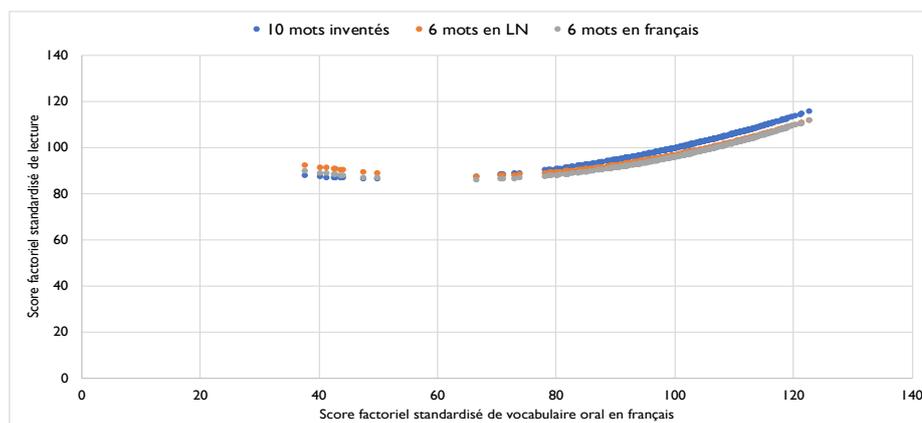


Il est possible que le niveau de vocabulaire oral en français des élèves intervienne de façon différente sur leur niveau de lecture selon la nature et la langue des mots qu'ils doivent lire. Pour examiner cette question, nous avons construit des modèles de régression, en prenant comme variables à expliquer les scores de lecture des élèves (i) sur les 10 mots inventés, (ii) sur les 6 mots en LN et (iii) sur les 6 mots en français. Les résultats des modèles sont fournis dans le tableau 51 ci-dessous, et sont complétés par le graphique 18 ci-après, qui représente la relation entre le score de vocabulaire oral et les trois scores de lecture. Les analyses indiquent principalement et de façon attendue, que le niveau de vocabulaire oral en français explique plus fortement la variabilité des scores en lecture de mots français qu'en lecture de mots inventés ou issus des LN.

Tableau 51 : Analyses de régression du score de lecture en CEI selon le niveau de vocabulaire oral en français (en tenant en compte des caractéristiques des élèves et de leur environnement familial)

Variable	Score sur les 10 mots inventés	Score sur les 6 mots en LN	Score sur les 6 mots en français
Constante	106,90 ***	116,82 ***	112,06 ***
Score factoriel de vocabulaire	-0,64 ***	-0,90 ***	-0,87 ***
Score factoriel de vocabulaire ²	0,01 ***	0,01 ***	0,01 ***
R ²	17,9%	16,8%	20,7%

Graphique 18 : Relation entre le score de vocabulaire oral en français et les scores aux différentes parties du test de lecture



En résumé, quelle que soit la classe, la relation entre le niveau de vocabulaire oral en français des élèves et leur niveau de prélecture ou de lecture est positive et significative. De plus, toujours quelle que soit la classe, un certain niveau de vocabulaire oral en français doit être atteint pour qu'il ait un impact positif sur les niveaux de prélecture ou de lecture. Alors que le niveau de vocabulaire oral en français de quasiment tous les élèves à l'entrée au CI se situe en dessous de ce seuil, c'est encore le cas pour 41% des élèves de CP mais pour seulement 7% de ceux de CEI. Le CEI semble donc être la période la plus propice pour introduire l'apprentissage de la lecture en français L2, tout au moins au niveau du mot. Cette conclusion est renforcée par le fait que les évaluations effectuées à cette époque ont également permis de relever que le niveau de vocabulaire oral en français explique plus fortement la variabilité des scores en lecture de mots français comparativement à des mots inventés ou issus des LN.

Jusqu'ici, nous avons considéré l'effet du niveau de vocabulaire oral en français des élèves sur leur niveau de prélecture ou de lecture de façon séparée pour chaque niveau. Dans la mesure où l'on dispose des scores obtenus à quatre items communs aux trois niveaux (*sa, la, dé, ti*), il est possible d'estimer si l'effet du niveau de vocabulaire oral en français sur le niveau de décodage est différent selon les classes et avec quelle ampleur. Comme précédemment, on pose l'hypothèse (i) que les élèves doivent avoir atteint un certain seuil de vocabulaire oral en français avant d'aborder l'apprentissage de la lecture et (ii) que ce seuil se situe avant que l'effet du niveau de vocabulaire oral en français sur le niveau de prélecture soit le plus fort.

De façon concrète, on cherche donc à expliquer le score de prélecture aux quatre items communs (SCLECT) des élèves avec :

$$\text{SCLECT} = f(\text{CIFAM}, \text{CP}, \text{CEI}, \ln(\text{SCVOCA}), \text{CPIInSCVOCA}, \text{CEIInSCVOCA})$$

- . Leur niveau (CI comme modalité de référence ou modalité omise, CP et CEI) ;
- . Le score de vocabulaire oral en français qu'ils ont obtenu ($\ln(\text{SCVOCA})$) ; on utilise ici le logarithme népérien⁴³ (une note pour expliquer) du score de vocabulaire oral en français, pour tenir compte de la relation non linéaire entre le niveau de vocabulaire et le niveau de lecture ou prélecture, telle que nous l'avons vu plus haut) ;
- . Deux variables d'interaction (CPIInSCVOCA et CEIInSCVOCA) entre le niveau et le logarithme népérien du score de vocabulaire oral en français (CPIInSCVOCA = CP x $\ln(\text{SCVOCA})$ et CEIInSCVOCA = CEI x $\ln(\text{SCVOCA})$) ;
- . En tenant compte de leurs caractéristiques et de celles de l'environnement familial (CIFAM).

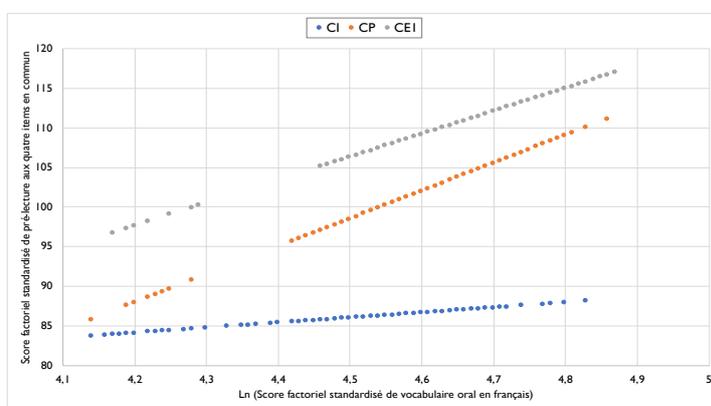
L'interprétation se fait alors selon le degré de significativité et la valeur de chacun des coefficients attachés aux deux variables d'interaction, tels que fournis par les estimations du modèle de régression. Il ressort du modèle que les coefficients attachés aux variables d'interaction sont toutes deux significatives et positives : elles s'élèvent à 29 pour celle caractérisant le CP et à 23 pour celle caractérisant le CEI. Pour davantage de lisibilité, nous avons simulé, sur la base des coefficients du modèle de régression, les scores qu'obtiendraient

⁴³. Il est possible que la relation entre la variable à expliquer (ici, le score de pré-lecture) et la variable explicative (ici, le score de vocabulaire) ne soit pas linéaire. Une indication en est que la représentation graphique de la relation entre les deux variables n'est pas une droite mais une courbe. Dans la mesure où la méthode des moindres carrés ordinaires (utilisée ici) est une méthode de régression linéaire, qui présume que les relations modélisées sont linéaires, il est alors nécessaire de transformer la variable explicative. Plusieurs transformations non linéaires sont possibles, celle utilisée ici est la transformation logarithmique. Elle s'avère plus simple à interpréter ici, dans la mesure où les analyses prennent en compte un effet d'interaction entre deux variables explicatives.

les élèves selon leur classe et leur niveau de vocabulaire oral en français. Les résultats sont présentés dans le graphique 19 ci-après.

La pente des trois courbes est ascendante, ce qui indique un effet positif et significatif du niveau de vocabulaire oral en français sur leur niveau de décodage dans les trois classes. Le degré d'ascendance des pentes des courbes est cependant variable selon les classes. Elle est modérée à l'entrée au CI : l'effet du niveau de vocabulaire oral en français sur leur niveau de décodage est alors le plus faible. A l'inverse, c'est à l'entrée au CP que la pente de la courbe est la plus ascendante : c'est à ce niveau que l'effet du niveau de vocabulaire oral en français sur leur niveau de décodage est le plus fort. Il est donc crucial que les élèves disposent à l'entrée au CP d'un certain niveau de vocabulaire oral en français (niveau que nous avons mis en lumière dans les analyses précédentes) pour « maximiser » leur apprentissage de la lecture. En CEI, la relation entre le niveau de vocabulaire oral en français des élèves et leur niveau de prélecture est également forte, mais moins qu'à l'entrée au CP.

Graphique 19 : Scores selon la classe et le niveau de vocabulaire oral en français



Ces analyses permettent de cibler le CI comme étant l'année où les élèves doivent améliorer leur niveau de vocabulaire oral en français (parce qu'à l'issue de cette année, une fois à l'entrée au CP, un meilleur niveau de vocabulaire oral en français leur permettra d'atteindre un meilleur niveau de prélecture). Nous reviendrons sur cette question dans la discussion générale.

Toujours dans la perspective d'expliquer la variabilité des niveaux de prélecture et lecture des élèves, il importe de mobiliser des facteurs liés à leur expérience scolaire l'année précédant l'enquête. Pour cela, nous avons construit des modèles de régression dans lesquels la variable à expliquer est le score obtenu en prélecture pour les élèves de CP et en lecture pour les élèves de CEI, et les variables explicatives sont, en plus des caractéristiques des élèves, de leur environnement familial et de leur score en vocabulaire oral en français déjà examinées plus haut (et sur lesquelles nous ne reviendrons pas), les modalités d'organisation des classes et les activités mises en œuvre par les enseignants l'année précédant l'enquête. Les estimations pour les élèves de CP sont présentées dans le tableau 52 ci-après.

Tableau 52 : Analyses de régression du niveau de prélecture des élèves de CP selon certaines variables sociales et les conditions et activités d'enseignement-apprentissage en 2017-2018

(avec prise en compte des caractéristiques des élèves et de leur environnement familial, ainsi que de leur score de vocabulaire oral en français)

Variable		Modalités	CP
Constante			110,34 ***
Caractéristiques de la classe	Taille de la classe		-0,12 **
	Programme LPT	Non Oui	réf. 6,89 ***
Caractéristiques de l'enseignant	Sexe	Femme Homme	réf. -6,01 ***
Activités pédagogiques	Pôle Expression et compréhension orale		-3,02 ***
R ²			28,3%

Parmi les variables liées à l'expérience scolaire des élèves, qui concourent à l'explication de la variabilité de leurs performances au test, il ressort d'abord que la taille de la classe a un effet négatif sur le niveau de prélecture des élèves. Cet effet est cependant plutôt modéré, un élève supplémentaire dans la classe correspondant à une baisse du score de prélecture des élèves de 0,12 point. De plus, les élèves de CP qui étaient dans une classe de CI participant au programme Lecture Pour Tous l'année qui a précédé l'enquête, et qui ont donc bénéficié d'heures d'enseignement de lecture en langue nationale, ont un niveau de prélecture supérieur (d'un peu moins de 7 points) à celui des élèves qui n'ont pas bénéficié de ce dispositif. Parmi les caractéristiques des enseignants, seul leur genre introduit des différences : les élèves qui avaient une enseignante plutôt qu'en enseignant en CI ont un niveau de prélecture un peu meilleur (d'un peu plus de 6 points). Sur le plan des activités mises en place par les enseignants, les analyses indiquent une relation négative entre une plus grande fréquence des activités dédiées à l'expression et à la compréhension orale et le niveau de prélecture des élèves.

Concernant le CEI (tableau 53 ci-après), la part de variance expliquée la plus élevée est observée sur la lecture des mots issus du français (33% contre 26% pour les mots inventés et ceux issus des LN). Le niveau de lecture des élèves s'améliore lorsqu'ils disposaient et utilisaient un livre d'exercices de français (même partagé) en classe de CP. Il augmente également avec le degré d'aisance en français de leur enseignant : les scores varient de 99 lorsque leur enseignant de CP s'estime peu à l'aise en français (degré de 0) à 106 lorsqu'il se considère très à l'aise en français (degré de 10). Concernant les activités mises en place par les enseignants, le niveau en lecture de mots inventés (et seulement celui-là) s'améliore avec la pratique d'exercices dédiés au code. En revanche, le temps passé à travailler sur la langue orale, qui n'a aucun effet sur la lecture de mots inventés, a un impact positif plus marqué sur les mots issus du français que sur ceux des LN.

Tableau 53 : Analyses de régression du score de lecture des élèves de CEI selon certaines variables sociales et les conditions et activités d'enseignement en 2017-2018 (avec prise en compte de leurs caractéristiques et de leur environnement familial, ainsi que de leur score de vocabulaire oral en français)

Variable		Modalités	Score sur les 10 mots inventés	Score sur les 6 mots en LN	Score sur les 6 mots en français	Score global de lecture
Constante			108,10 ***	119,00 ***	205,62 ***	110,17 ***
Caractéristiques de la classe	Taille de la classe		-0,12 **	-	-	-
	Livre d'exercic	Non	réf.	réf.	réf.	réf.
		Oui (même partagé)	4,27 ***	7,10 ***	5,81 ***	6,29 ***
	Durée de l'année (en semaine)		-	-	-5,91 *	-
Durée de l'année ² (en semaine)		-	-	0,09 *	-	
Caractéristiques de l'enseignant		Degré d'aisance en français	0,78 ***	-	1,14 ***	0,86 ***
Activités d'enseignement-apprentissage	Pôle Enseignement de la langue		-	0,02 ***	0,03 ***	0,02 ***
	Pôle Apprentissage du code		1,59 *	-	-	-
	Pôle Récit par cœur		-	-	-2,9**	-2,11 *
	Pôle Conjugaison		-3,69 ***	-2,63 ***	-1,84 *	-2,25 ***
R ²			26,1%	25,9%	32,7%	31,8%

En résumé, concernant les variables liées à l'expérience scolaire des élèves, un des résultats les plus intéressants est que ceux de CP qui ont participé au programme Lecture Pour Tous l'année qui a précédé l'enquête, et qui ont donc bénéficié d'heures d'enseignement de lecture en langue nationale, ont un niveau de prélecture supérieur (environ ½ écart-type) à celui des élèves qui n'ont pas bénéficié de ce dispositif. Au même niveau scolaire, seul le genre des enseignants a un effet positif (les élèves qui avaient une enseignante en CI ont un niveau de prélecture supérieur d'environ ½ écart-type). En revanche, une relation négative est relevée entre une plus grande fréquence des activités dédiées à l'expression et à la compréhension orale et le niveau de prélecture des élèves.

Le principal facteur qui a un impact positif sur leur niveau de lecture de mots (inventés et fréquents) des élèves vus en CE1 est l'utilisation d'un livre d'exercices de français (même partagé) en classe de CP. Quant aux activités mises en place par les enseignants, la pratique d'exercices dédiés au code a un impact positif sur le niveau en lecture de mots inventés (et seulement celui-là) alors que le temps passé à travailler sur la langue orale, qui n'a aucun impact sur la lecture de ces items, a un effet positif sur la lecture des mots en français.

4. SYNTHÈSE ET DISCUSSION

4.1. Rappel des hypothèses

Lorsqu'ils doivent apprendre à lire dans leur L1, la plupart des élèves ont, dans cette langue, un niveau de langage oral qui leur permet de comprendre le langage écrit. Ce n'est souvent pas le cas de ceux qui apprennent à lire dans une L2. Notre première hypothèse était donc que, pour comprendre ce qu'ils lisent en français L2, les élèves sénégalais devraient non seulement être capables de comprendre à l'oral, dans cette langue, des mots isolés mais aussi des suites de mots. Notre seconde hypothèse comprenait trois volets, l'un ($H2^{\text{voc-lec}}$) relié à la première hypothèse. Le niveau du vocabulaire oral des élèves en français L2 devrait augmenter avec leur niveau de scolarité ($H2^{\text{voc}}$), tout comme leur niveau de prélecture ou de lecture ($H2^{\text{lec1}}$). Ces évolutions devraient permettre de cerner à partir de quel degré de maîtrise de la langue orale (mots isolés et groupes de mots) il devient possible d'apprendre à lire dans cette L2 ($H2^{\text{voc-lec}}$).

Toutefois, pour comprendre un texte lu dans une écriture alphabétique, il faut, d'une part, avoir un bon niveau de compréhension de la langue orale dans laquelle s'effectue la lecture et, d'autre part, avoir automatisé le décodage, la rapidité de cette automatisation dépendant du degré de régularité des relations graphème-phonème dans la langue cible. De plus, la maîtrise du décodage étant un puissant mécanisme d'auto-apprentissage (ce qui provient du fait que, quand les relations graphème-phonème les plus régulières ont été acquises, il devient plus facile d'apprendre celles qui le sont moins), apprendre d'abord à lire dans une des LN des enfants devrait faciliter cet apprentissage en français L2. En effet, d'une part, l'orthographe de ces LN (wolof, pulaar et seereer, dans notre projet) est plus régulière que celle du français, ce qui facilite la compréhension des relations graphème-phonème et le transfert à une orthographe moins régulière. D'autre part, les enfants maîtrisent mieux leur L1 à l'oral que le français. Les résultats obtenus en CEI dans l'épreuve de lecture qui contient des items permettant d'examiner la maîtrise du décodage (mots inventés) et celle de la procédure lexicale d'identification des mots écrits (mots issus du français ou des LN) permet d'évaluer, à ce niveau scolaire, les prédictions précédentes. Notre troisième série d'hypothèse concernait donc les relations entre le niveau de vocabulaire oral en français et celui des procédures spécifiques à la lecture (décodage et procédure lexicale d'identification des mots écrits). Si les élèves de CEI ont recours à la procédure lexicale d'identification des mots écrits, les mots de leur LN devraient être mieux lus que les mots inventés ou ceux en français ($H3^{\text{lec}}$). Toutefois, le niveau de vocabulaire oral en français devrait rendre compte des scores en lecture de mots dans cette langue ($H3^{\text{voc-lec}}$). De plus, les activités scolaires en français autour du vocabulaire et de la compréhension orale devraient avoir surtout un impact sur le vocabulaire oral en français ($H4^{\text{voc}}$) alors que l'impact de celles autour du code devrait se manifester sur le décodage ($H4^{\text{lec}}$). Il en est de même de la participation au programme Lecture Pour Tous ($H4^{\text{lec1}}$) qui, par ricochet, devrait avoir un impact positif sur la procédure lexicale d'identification des mots écrits en français ($H4^{\text{lec2}}$) ainsi que sur le niveau de vocabulaire oral dans cette langue ($H4^{\text{lec-voc}}$).

D'autres facteurs en lien avec l'environnement des élèves sont susceptibles d'influencer leurs niveaux de vocabulaire oral en français, et de prélecture ou lecture. Certains concernent leur milieu familial ($H5$) : les niveaux de vocabulaire oral en français ($H5^{\text{voc1}}$) et de lecture ($H5^{\text{lec1}}$) sont supposés être meilleurs chez les élèves issus d'un milieu favorisé ; de même, ceux qui bénéficient chez eux d'une personne qui passe du temps à leur parler, leur apprendre des mots nouveaux, ou leur lire des histoires en français, devraient avoir de meilleurs scores de

vocabulaire oral en français ($H5^{voc2}$) et de lecture ($H5^{lec2}$), que ceux qui vivent dans un milieu moins favorisé. D'autres facteurs sont en lien avec l'école ($H6$), en particulier, la place accordée aux activités autour du vocabulaire oral en français et à celles sur le décodage. Sont également supposées avoir un effet facilitateur sur l'acquisition du vocabulaire oral en français ($H6^{voc}$) et sur la lecture dans cette langue ($H6^{lec}$), les compétences des enseignants en français, appréhendées par leur niveau de formation ($H6^{voc1}$ & $H6^{lec1}$) ou la perception qu'ils en ont ($H6^{voc2}$ & $H6^{lec2}$).

4.2. Evaluation des hypothèses sur le niveau du langage oral en français et de lecture

4.2.1. Épreuve de vocabulaire (du mot au groupe de mots)

L'hypothèse H1 stipulant que pour pouvoir comprendre ce qu'ils lisent en français, les élèves devraient non seulement être capables de comprendre à l'oral des mots isolés, mais aussi des suites de mots, n'est pas corroborée. En effet, dans l'épreuve dans laquelle il fallait effectuer des actions simples (en réponse à une question portant sur un mot), puis des actions plus complexes (en réponse à une question impliquant plusieurs mots), seuls 4% des élèves de CI, 10% de ceux de CP et 16% de ceux de CEI se sont avérés capables d'aller au-delà du critère d'arrêt fixé après les actions simples. **La plupart des élèves évalués n'ont donc qu'une compréhension du français oral limitée au mot.**

En revanche, le premier volet de la seconde hypothèse ($H2^a$) est corroboré. En effet, le score moyen de vocabulaire oral en français des élèves (qui a une moyenne de 100 et un écart-type de 15) augmente de presque un écart-type entre le CI et le CP et d'environ $\frac{1}{2}$ écart-type entre le CP et le CEI. **C'est donc au cours de la première année du primaire que le niveau de vocabulaire oral en français augmente le plus.**

En outre, si le niveau de vocabulaire oral en français des élèves est assez peu dépendant du contexte dans lequel ils évoluent (les parts de variance expliquées sont faibles et elles baissent avec les niveaux scolaires : 27% en CI, 21% en CP et 10% en CEI ; cf. tableau 28), en conformité avec l'hypothèse $H5^{a1}$, **le poids du milieu socioculturel est significatif aux trois niveaux scolaires** (5 points en plus pour les élèves de milieu favorisé en CP et respectivement 3 et 4 en CI et CEI). **Il en est de même du degré de proactivité pour lire, parler et apprendre des mots en français dans le milieu familial** : quand ces trois activités sont pratiquées au sein de la famille, le niveau de vocabulaire augmente de 6 points en CI et CEI et de 4 en CP. En lien avec ce résultat, qui corrobore l'hypothèse $H5^{a1}$, regarder la télévision en français améliore le niveau de vocabulaire en CI (6 points) et CP (4 points), mais pas en CEI. Le seul autre facteur ayant un effet sur le niveau de vocabulaire oral en français est l'appartenance à un milieu urbain (9 points en plus au CI et CP et 8 au CEI).

Enfin, contrairement à l'hypothèse $H4^{voc}$, **les activités pédagogiques dans le domaine du vocabulaire et plus largement dans celui de la compréhension du langage, n'ont pas d'effet positif sur le niveau de vocabulaire oral en français.** Il ressort même des analyses que les élèves de CEI ont un meilleur niveau de vocabulaire oral en français lorsqu'en CP leurs maîtres ont passé plus de temps à travailler sur l'enseignement de la grammaire, et moins de temps à travailler sur la compréhension en lecture (cf. le tableau 46). En outre, si le temps

dédié à ces deux domaines à un effet significatif (positif ou négatif) sur le niveau de vocabulaire des élèves, le contenu des activités entrant dans ces deux domaines n'en a pas⁴⁴. La question de l'impact sur le vocabulaire des activités scolaires dont les élèves ont bénéficié est réexaminée dans la section 4.3.2.

4.2.2. Épreuve de lecture

Le second volet de la deuxième hypothèse (H2^{lec}) est corroboré. En effet, d'après les résultats d'une épreuve de décodage commune aux trois classes, l'amélioration est de 16 points (un peu plus d'un écart-type) entre le CI et le CP et de 10 points entre le CP et le CEI. **C'est donc au cours de la première année du primaire que le niveau de décodage augmente le plus.**

Quant à l'hypothèse H3^{lecl} à propos de l'utilisation de la procédure lexicale d'identification des mots écrits, elle n'est pas corroborée : la différence entre les scores en lecture de mots inventés et de mots en LN de même longueur n'est pas significative. Cette différence est, en revanche, significative entre ces deux scores et celui sur les mots en français, les élèves lisant en moyenne un peu mieux les mots inventés et les mots de leur LN que les mots en français. L'utilisation d'une même procédure pour lire les mots inventés et les mots des LN ressort aussi de l'analyse des corrélations : celle entre ces deux sous-tests de lecture est très élevée (0,84). **Ces résultats suggèrent que les élèves, à ce niveau scolaire (CEI), utilisent largement une même procédure de lecture pour identifier les mots, à savoir le décodage.**

Les caractéristiques des élèves et de leur environnement familial n'ont que peu d'emprise sur leur niveau de prélecture (tableaux 46 et 47). Toutefois, en conformité avec l'hypothèse H5^{lecl}, ceux venant des familles moyennes et favorisées ont un meilleur niveau de lecture que les autres mais uniquement en CI (environ 6 points) et en CP (3 points), pas en CEI. L'effet des pratiques de lecture à la maison est, quant à lui, toujours positif (6 points en CI et CP, 5 en CEI), comme attendu (cf. H5^{lec2}). Ces résultats, qui indiquent que les élèves dont les parents sont plus éduqués et / ou disposent des meilleures ressources éducatives obtiennent de meilleurs scores en lecture, sont conformes aux données de la littérature⁴⁵.

Un seul résultat permet de valider l'hypothèse H4^{lec} à propos des effets des activités scolaires autour du code sur les résultats en décodage. En effet, les activités dans ce domaine n'ont, en CP, aucun effet sur la maîtrise du décodage évaluée par la lecture de mots inventés (tableau 52). En revanche, elles ont un impact sélectif sur ce type d'items en CEI, mais pas sur la lecture de mots issus des LN ou du français (tableau 53). Le niveau de lecture s'améliore aussi chez les élèves de CEI lorsqu'ils disposaient et utilisaient un livre d'exercice en français en classe de CP, et cet effet se manifeste aussi bien sur les mots inventés que sur ceux issus des LN ou du français. Il s'améliore aussi, mais de façon sélective sur les mots français quand leurs enseignants se sentent à l'aise dans cette langue (H6^{lec2}).

⁴⁴. D'autres résultats surprenants ont été relevés : la préscolarisation n'a un impact positif que sur les scores en CP (+ 8 points), pas sur ceux des élèves de CI ou de CEI ; le genre des élèves est au profit des garçons (+ 3 points).

⁴⁵. Voir l'étude de Braga, Checchi, Garrouste et al. (2019), basée sur les données de PIRLS.

Un des résultats observés en CP⁴⁶ permet de valider l'hypothèse H4^{lec}. Ainsi, **les élèves qui étaient dans une classe de CI participant au programme Lecture Pour Tous l'année qui a précédé celle durant laquelle ont eu lieu les observations, et qui ont donc bénéficié d'heures d'enseignement de lecture en langue nationale, ont un niveau de prélecture supérieur à celui des élèves qui n'ont pas bénéficié de ce dispositif** (environ ½ écart-type).

En dehors de l'hypothèse H3^{lec} concernant l'utilisation de la procédure lexicale d'identification des mots écrits et de celle sur le rôle des activités autour du décodage sur les compétences dans ce domaine (H4^{lec}), tous les effets observés sont soit conformes aux attentes (la possibilité d'utiliser des livres d'exercices en français), soit à nos hypothèses. C'est en particulier le cas pour les hypothèses sur l'incidence du milieu socioculturel et des activités autour du français à la maison (H5^{voc} et H5^{lec}) et pour celle sur la perception qu'ont les enseignants de leur degré de maîtrise du français sur le niveau de lecture de leurs élèves dans cette langue (H6^{lec}).

L'hypothèse sur l'effet des exercices autour du code sur les résultats en décodage (lecture de mots inventés) est réexaminée dans la section suivante (4.3.2).

4.3. Autres facteurs ayant une incidence sur les capacités évaluées

Parmi les facteurs supposés avoir un impact sur le niveau de vocabulaire oral en français ou sur celui de lecture, certains sont source de résultats contradictoires (ceux sur les effets de la localisation géographique), d'autres sont contrintuitifs (par exemple, ceux sur les effets négatifs des diplômes universitaires), d'autres encore, qui vont dans le sens attendu, ne se retrouvent pas de façon systématique aux différents niveaux scolaires (par exemple, ceux sur l'utilisation de livres d'exercices). La suite de cette section examine ces différents problèmes.

4.3.1. Résultats contradictoires à propos de l'environnement géographique

Parmi les facteurs supposés avoir un impact sur le niveau de vocabulaire oral en français qui ont été examinés dans les trois classes, la localisation géographique est celui qui introduit le plus de différence : les élèves résidant en milieu urbain obtiennent en moyenne un score supérieur de 8 à 9 points (plus de la moitié d'un écart-type) par rapport à ceux habitant en milieu rural. A l'inverse, le niveau de prélecture des élèves s'avère meilleur parmi ceux résidant en milieu rural, mais uniquement en CI, pas en CP. Le bénéfice est d'environ 5 points, l'écart entre les deux groupes atteignant même près de 9 points lorsqu'ils sont appariés sur leur niveau de vocabulaire oral en français. Il est possible d'expliquer l'effet positif de l'urbanité sur le niveau de vocabulaire oral en français par la plus forte proportion de personnes parlant le français en milieu urbain qu'en milieu rural⁴⁷. En revanche, l'effet bénéfique, sur le niveau de décodage, de l'appartenance à un milieu rural, est difficilement interprétable.

⁴⁶. Seuls les élèves de ce niveau ont pu bénéficier du programme Lecture Pour Tous, pas ceux de CE1.

⁴⁷. La proportion des parents qui se déclarent à l'aise en français est plus forte en milieu urbain qu'en milieu rural.

4.3.2. A propos des activités scolaires

Pour rappel, les années au cours desquelles ont été effectuées les activités ciblées dans cette section sont celles qui ont précédé les évaluations (CI pour celles de CP et CP pour celles de CEI).

4.3.2.1. Activités de vocabulaire

Les activités sur le français dans le milieu familial ont un effet sur le niveau de vocabulaire oral des enfants, quelle que soit la classe. Ce n'est, de façon très surprenante, pas le cas des activités pédagogiques autour du vocabulaire dont les élèves ont bénéficié l'année qui a précédé celle durant laquelle ont eu lieu les observations. Il ressort même des analyses que les élèves de CEI ont un meilleur niveau de vocabulaire oral en français lorsqu'en CP leurs maîtres ont passé plus de temps à travailler sur l'enseignement de la phrase et de la grammaire, et moins de temps à travailler sur la compréhension en lecture (cf. le tableau 46). En outre, si le temps dédié à ces deux domaines a un effet significatif (positif ou négatif) sur le niveau de vocabulaire des élèves, le contenu des activités entrant dans ces deux domaines n'en a pas.

Ces résultats sont étranges à double titre. En effet, il est évident qu'un travail sur la compréhension en lecture doit avoir une incidence sur le vocabulaire. En témoignent les travaux de recherche sur les liens entre vocabulaire oral et compréhension de la langue orale ou écrite⁴⁸. En revanche, les liens entre vocabulaire et grammaire sont peu documentés. Ces résultats suggèrent que, plus que le temps passé à exercer certaines activités, c'est la qualité des exercices proposés qui compte. Il est fort possible que les exercices proposés par les maîtres dans ce domaine ne soient pas les plus appropriés. Quelques suggestions, reprises du cadre de référence, sont présentées à la fin du rapport.

4.3.2.2. Activités autour du code

Un seul résultat permet de valider l'hypothèse H4^{lec} à propos des effets des activités scolaires autour du code sur les résultats en décodage : celui obtenu en CEI. **A la différence du CEI, en effet, les activités sur le code n'ont, en CP, aucun effet sur la maîtrise du décodage évaluée par la lecture de mots inventés** (tableau 52). En revanche, elles ont un impact sélectif sur ce type d'items en CEI, mais pas sur la lecture de mots issus des LN ou du français (tableau 53). Les activités autour du code sont pourtant connues pour avoir de façon systématique un effet positif sur le décodage, en LI, comme en L2⁴⁹.

Il est possible d'expliquer ce résultat surprenant par le bruit introduit par un travail trop systématique sur le nom des lettres en CI. En effet, tous les enseignants de ce niveau scolaire effectuent des activités dans ce domaine, tous les jours pour une vaste majorité (69%) alors qu'ils sont un peu moins nombreux en CP à pratiquer journalièrement des exercices de ce type (60%). Si, pour mettre en relation les lettres de l'alphabet avec les

⁴⁸. Voir Hoover & Gough (1990); Perfetti (2007); Perfetti & Stafura (2014); Quinn et al. (2015); Tunmer & Chapman (2012).

⁴⁹. Voir pour des synthèses en LI : Castles et al. (2018) et en L2 : August & Shanahan (2006) et Murphy (2018); voir aussi le cadre de référence.

phonèmes qui leur correspondent, il est nécessaire de connaître le son de ces lettres, utiliser le nom des lettres peut gêner l'apprentissage des relations graphème-phonème et phonème-graphème pour les consonnes. En effet, une consonne doit toujours *sonner* avec un autre élément, à savoir une voyelle : en conséquence, le nom des consonnes de l'alphabet est accompagné par une voyelle : <é>, dans *bé, dé, pé, él, ér...* ; <i> dans *ji* ; <a> dans *ka* ; <u> dans *ku...* Un élève qui utilise cette connaissance peut donc lire « pays » pour <pi> et « déesse » pour <ds> (ou écrire ces deux mots *pi* et *ds*). Ce problème, signalé dans la grammaire de Port-Royal⁵⁰, a fait l'objet d'une étude française qui a montré que, si la connaissance du nom des lettres pour les voyelles⁵¹ à 5 ans est un des prédicteurs du futur niveau de décodage trois ans plus tard (en fin de CEI), ce n'est pas le cas pour le nom des consonnes⁵². C'est ce qui explique pourquoi, alors que la connaissance de l'alphabet était évaluée dans la première version d'EGRA par une épreuve d'identification du nom des lettres, elle l'est maintenant par une épreuve portant sur le son des lettres, cette connaissance étant directement liée à la capacité de décodage⁵³.

4.3.2.3. Une vue d'ensemble des effets des activités scolaires sur les compétences évaluées

Dans l'ensemble, l'absence d'impact des activités d'enseignement mises en œuvre sur le niveau de vocabulaire oral en français des élèves suggère que les questions posées aux enseignants sur ces aspects ne sont probablement pas assez précises pour rendre compte de ce qui se passe en classe. Des observations dans la classe auraient sans doute permis de mieux appréhender les durées et les contenus des activités pédagogiques et d'analyser plus finement leur effet sur le niveau de vocabulaire oral en français des élèves, ainsi que sur leur capacité de décodage⁵⁴. Toutefois, pour être valides, des observations de ce type doivent être guidées par un cadre de référence solide, tel que celui qui est joint au présent rapport, dans lequel sont intégrées des suggestions d'activités autour du vocabulaire et une proposition de progression pédagogique pour l'apprentissage de la lecture.

4.3.3. A propos du niveau de diplôme des enseignants⁵⁵

Les données à ce propos sont peu robustes. En effet, les résultats des élèves diffèrent en fonction du type de diplôme de leurs enseignants (diplôme universitaire ou professionnel) et du niveau dans lequel ils ont enseigné (CI ou CP). Ainsi, seuls les élèves de CEI qui ont eu en CP un enseignant ayant le CAP ont un meilleur niveau de vocabulaire (de l'ordre de 5 points) que

⁵⁰. Arnauld & Lancelot (1660).

⁵¹. Ce qui est valable pour le français ne l'est pas pour toutes les langues, en particulier l'anglais, langue dans laquelle les voyelles sont peu stables (elles sont souvent diphtonguées : cf. la prononciation de « do », voir Delattre, 1965). Toutefois, en anglais, il a été montré qu'il est plus facile de reconnaître que *beech* par rapport à *bone* commence par la lettre (qui se prononce /bi/ en anglais) ou *deaf* par rapport à *loaf* se termine la lettre <f> (qui se prononce comme en français) : voir Treiman, Tincoff & Richmond-Welty (1996) ; voir aussi Treiman & Tincoff (1997).

⁵². Voir Piquard-Kipffer & Sprenger-Charolles (2013).

⁵³. RTI International (2016).

⁵⁴. Voir Bressoux (2001).

⁵⁵. Pour rappel, le diplôme (universitaire ou professionnel) est celui des enseignants que les élèves ont eu l'année qui a précédé les évaluations (CI pour celles de CP et CP pour celles de CEI).

les autres (ceux qui ont eu en CP un enseignant sans formation pédagogique ou titulaire du CEAP). Ce n'est pas le cas pour les élèves de CP. A l'inverse, la possession d'un diplôme universitaire a un effet négatif sur le vocabulaire oral des élèves en français, mais encore une fois uniquement à un niveau scolaire : seuls les élèves de CP ayant eu un enseignant de CI titulaire d'une licence ou d'une maîtrise ont un niveau de vocabulaire oral en français inférieur à celui des élèves qui ont eu un enseignant moins diplômé. Ce dernier résultat, qui est fréquent dans la littérature sur l'évaluation des acquis des élèves en Afrique⁵⁶, est généralement expliqué par le fait que les enseignants ayant obtenu un diplôme universitaire peuvent avoir des aspirations professionnelles plus importantes que ceux d'un niveau inférieur, et de ce fait, être moins motivés. De plus, le niveau du diplôme universitaire n'est pas un gage de la maîtrise des connaissances de base qui doivent être enseignées à l'école primaire. Il faudrait en effet s'assurer de l'adéquation entre le contenu de leur diplôme et ces connaissances.

4.3.4. A propos des modalités d'organisation de la classe

Il faut rappeler que ces modalités correspondent à la classe dans laquelle les élèves étaient scolarisés l'année qui a précédé les évaluations (CI pour celles de CP ; CP pour celles de CEI). L'augmentation de la taille de la classe a un effet négatif sur le niveau de vocabulaire oral en français des élèves de CP et de CEI et sur le niveau de prélecture de ceux de CP (mais pas sur le niveau de lecture des élèves de CEI). Si cet effet est en général plutôt limité, il faut néanmoins souligner qu'il existe une forte variabilité des classes sur ce plan, le nombre d'élèves par classe variant de moins de 10 à plus de 100. En ce qui concerne le temps scolaire, une année de CI se déroulant sur une plage allant de 32 à 36 semaines serait optimale, alors qu'une durée optimale pour une année de CP serait d'environ 25 semaines. De plus, l'interruption des activités scolaires pendant plus d'un mois pendant l'année de CI a un effet négatif sur le niveau de vocabulaire oral en français en CP (d'environ 8 points), qui n'existe pas pour les élèves de CEI. En revanche, la durée de l'année scolaire n'a pas d'effet sur le niveau de lecture des élèves.

Ces résultats signalent que les élèves ont besoin d'un temps plus long et plus continu en CI qu'en CP pour améliorer leur niveau de vocabulaire oral en français. Ils suggèrent aussi qu'il serait bien de limiter les effectifs des classes de CI, ce qui devrait avoir un impact positif sur le niveau de vocabulaire oral des élèves en français ainsi que leur niveau de lecture dans leur LN. Toutefois, il serait intéressant d'analyser plus finalement le temps effectif d'enseignement dont les élèves ont bénéficié, en prenant par exemple en compte le temps de fonctionnement de chaque école, la fréquence et la durée de l'absentéisme des enseignants et des élèves, ou encore le temps d'instruction consacré à chaque activité⁵⁷.

⁵⁶. Mingat & Suchaut (2000) ; Bernard, Tyab & Vianou (2004) ; Michaelowa & Wechtler (2006).

⁵⁷. Voir Abdazi (2007).

4.4. Implications des résultats de l'étude pour les politiques éducatives

4.4.1. Implications des résultats concernant les questions au centre de l'étude

4.4.1.1. A partir de quand peut-on commencer à apprendre à lire en français L2 ?

Un des objectifs principaux de la présente étude était d'évaluer le niveau de maîtrise du vocabulaire oral en français des élèves dans les trois premières années du primaire afin de définir le moment le plus approprié pour débiter l'enseignement de la lecture dans cette langue. Il a été relevé, d'une part, un pallier qui se situe en début du CEI. A cette époque, d'une part, le niveau atteint par les élèves en décodage (lecture de mots inventés ou de mots fréquents issus soit du français, soit de leur LN) augmente fortement quand leur niveau de vocabulaire oral en français a atteint un certain seuil (cf. graphique I6) et, d'autre part, la grande majorité des élèves (93%) a alors un niveau de vocabulaire oral en français suffisant, pour que cela se traduise de façon positive sur leur niveau de décodage. Il ressort aussi des analyses que le niveau de vocabulaire oral en français explique alors plus fortement les scores en lecture des mots issus du français que ceux pour les mots inventés ou issus d'une LN. Ces résultats permettent d'apporter une première réponse à notre hypothèse H2^c : **à partir du début du CEI, il est possible de commencer d'apprendre à lire des mots isolés en français.**

Toutefois, le niveau de vocabulaire des élèves de CEI n'est pas encore suffisant pour leur permettre de comprendre des textes écrits : presque 85% d'entre eux sont en effet alors incapables de comprendre une suite de mots français présentée à l'oral. Ce n'est bien entendu pas le cas pour des élèves qui ont le français comme LI : à l'entrée à l'école primaire (CP en France) presque tous sont capables de comprendre des énoncés oraux de 4 à 9 mots⁵⁸. Les résultats de la présente étude permettent de suggérer que **la compréhension de l'écrit en français L2 avec les élèves du Sénégal s'effectue, en début de CEI, en utilisant des textes en français lus par le maître.** C'est ce que proposent certains chercheurs concernant la lecture de textes en français LI au niveau du CP⁵⁹.

4.4.1.2. A partir de quand est-il possible de travailler sur le vocabulaire oral en français L2 ?

Les analyses permettent de cibler le CI comme étant la classe la plus propice pour travailler intensivement sur le vocabulaire oral en français. En effet, les résultats présentés dans la section D.3.4 signalent que les élèves ont besoin d'un temps plus long et plus continu en CI qu'en CP pour améliorer leur niveau de vocabulaire oral en français et que ce niveau s'améliore aussi quand l'effectif des classes de CI est inférieur à un certain seuil. C'est donc au CI qu'il faudrait intensifier les activités autour du vocabulaire oral en français, en ne les limitant pas à un travail sur des mots isolées (voir, pour des suggestions d'activités, la section D.4.2.1). Ce travail doit se prolonger au CP afin que le niveau de vocabulaire oral en français à l'entrée du CEI permette de commencer l'apprentissage de la lecture dans cette langue dans de bonnes conditions.

⁵⁸. Voir par exemple, Gentaz et al. (2013) ; Lecocq (1996).

⁵⁹. Par exemple, Fayol (2017).

4.4.2. Autres implications pour les politiques éducatives

Les deux parties suivantes intègrent quelques suggestions concernant ce qu'il faudrait présenter aux enseignants lors de leur formation (initiale et continue) afin de leur permettre d'intensifier, et de diversifier, les activités pédagogiques autour des deux piliers de l'apprentissage de la lecture : la compréhension orale (du mot isolé à la suite de mots) et les procédures spécifiques de lecture : la procédure lexicale d'identification des mots écrits et le décodage.

4.4.2.1. Compréhension orale et écrite

Les suggestions qui suivent portent aussi bien sur l'oral que sur l'écrit ainsi que sur des mots isolés ou des suites de mots pouvant être insérés dans une phrase. En effet, comme signalé dans le cadre de référence, d'une part, la compréhension du langage oral est largement similaire à celle du langage écrit⁶⁰. D'autre part, le travail sur la phrase ne se limite pas à la grammaire⁶¹.

L'étude a permis de constater que, alors que les activités autour de la langue française dans le milieu familial ont un effet à toutes les classes sur le niveau de vocabulaire oral des enfants⁶², ce n'est pas le cas pour les activités pédagogiques. Il serait donc souhaitable de présenter aux enseignants, lors de leur formation (initiale et continue), les différentes facettes du vocabulaire afin de leur permettre d'intensifier, et de diversifier, les exercices autour du vocabulaire.

Comme expliqué dans le cadre de référence, les mots sont définis par leur appartenance à une catégorie grammaticale (nom, verbe, pronom...) et à une catégorie sémantique (le *chien* est un animal). Au niveau sémantique, les relations peuvent être de plusieurs types : relations d'équivalence (synonymie : *auto* ou *voiture*) ou d'opposition (antonymie : *bon* versus *mauvais*, *début* versus *fin*). Elles peuvent aussi être fonctionnelles (le *marteau* sert à planter des clous) ou catégorielles et, dans ce cas, hiérarchique : hypéronymie (aller du particulier au général : du chat à l'animal) ou hyponymie (aller du général au plus particulier : le chat et un siamois).

Le sens des mots peut, en outre, être littéral (le *bouchon* de la bouteille) ou dérivé (le *bouchon* qui bouche la circulation). De plus, alors que certains mots n'ont qu'un seul sens, d'autres en ont plusieurs. Quelques-uns des mots « polysémiques » n'ont pas de relation sémantique : comme « avocat », qui est un fruit ou un métier, celui de la personne qui défend un accusé (et qui, pour cela, *parle*, cf. le radical *voc-*, qui se retrouve dans *vocabulaire* et *voyelle*). D'autres entretiennent des relations sémantiques : la *pâte* (à tarte) et *les pâtes* (à la bolognaise), à la différence des *pattes* du chien. Dans un dictionnaire, les mots de la première catégorie (les deux *avocats*) sont présentés dans deux entrées différentes alors que ceux de la seconde (les deux *pâtes*) sont le plus souvent dans une seule et même entrée.

Il existe aussi des unités lexicales complexes qui peuvent contenir deux noms (*bateau-mouche*), un nom plus un adjectif (*rouge gorge*) ou un verbe (*garde-fou*). Elles peuvent aussi intégrer une préposition entre deux noms (*pomme de terre*) ou deux verbes (*c'est-à-dire*). Ce qui est spécifique à la plupart de ces unités est que leur sens ne se réduit pas à celui de chacun des

⁶⁰. Voir par exemple, Gernsbacher, Varner & Faust (1990).

⁶¹. Voir par exemple, Lecocq (1996a) et Lecocq, Casalis, Leuwens et al. (1996b).

⁶². Comme cela l'a été montré dans d'autres études. Par exemple, d'après une étude dans laquelle le temps consacré chaque jour à la lecture hors de l'école par des élèves américains de 10-11 ans a été examiné, il est estimé que les lecteurs les plus faibles (10ème percentile) voient 60.000 mots par an alors que les meilleurs (90ème percentile) en voient plus de 4 millions (Anderson, Wilson & Fielding, 1988).

mots qui les composent : ainsi, le *garde-fou*, est un parapet qui empêche les gens de tomber, et non un individu devant garder les fous.

Un autre élément crucial dans un contexte bilingue est que l'étendue du vocabulaire dans certains domaines dépend de l'environnement. Ainsi, dans les langues d'Afrique, la probabilité que le mot *neige* soit fréquent est faible, tout comme l'est celle de trouver de nombreux mots pour spécifier la nature de la neige. En revanche, il peut y avoir dans ces langues de nombreux termes pour désigner la marche. Ainsi, en shona⁶³, environ 200 mots désignent les façons de marcher (*à reculons, avec une canne, à grandes enjambées, sur une longue distance, en soulevant la poussière, avec le dos vouté...*). **Les différences linguistiques liées à des facteurs culturels doivent être prises en compte dans un enseignement bilingue.** Il en est de même des expressions figées (comme *il pleut des cordes*)⁶⁴, qui diffèrent souvent fortement d'une langue à l'autre.

4.4.2.2. Les procédures spécifiques à la lecture : du décodage à la procédure lexicale

Alors que les résultats de l'étude indiquent que, en sensibilisant les enfants à la langue française et à la lecture dans cette langue dans le milieu familial, on peut améliorer leur niveau de prélecture, les activités autour du code à l'école n'ont qu'un effet très limité sur les résultats en lecture des enfants, cet effet n'ayant été relevé que sur la lecture de mots inventés chez les élèves de CEI. Ce résultat est très surprenant. En effet, on sait maintenant (cf. le bilan de plus de 30 ans de recherches)⁶⁵ que, aussi bien en L1 qu'en L2, un travail systématique, intensif et précoce sur ce type d'activité a un effet sur la maîtrise du décodage et, par voie de conséquence, sur l'automatisation de la procédure lexicale d'identification des mots écrits. Cette automatisation permet, en retour, au lecteur de consacrer ses ressources cognitives à la compréhension de ce qu'il lit.

Il serait donc souhaitable de présenter aux enseignants lors de leur formation (initiale et continue) les résultats des recherches sur les compétences qui doivent être développées chez les élèves pour leur faciliter l'apprentissage de la lecture dans une écriture alphabétique, et ce afin de rendre leur pratique pédagogique plus efficace. Le cadre de référence joint au rapport contient une suggestion de progression pour apprendre à lire en français. Cette progression tient compte de la fréquence des mots dans cette langue ainsi que de celle des correspondances graphème-phonème et de leur régularité. Elle tient également compte des principales spécificités de cette orthographe, la majeure partie étant liée à la morphologie⁶⁶, de nombreuses marques écrites étant soit absentes à l'oral (*ami / amie / amies...*) soit homophones (*chantais, chantait, chantaient...*)⁶⁷. Il faudrait aussi adapter cette progression aux spécificités des LN dans lesquelles les élèves du Sénégal vont commencer leur apprentissage de la lecture.

⁶³. Nancarrow (2004).

⁶⁴. Concernant les expressions figées du français, voir Caillies (2009).

⁶⁵. Voir par exemple les synthèses en anglais de Castles (2018) et de Clifton et al. (2016) ; en français, voir Dehaene et al. (2011) ; Sprenger-Charolles & Desrochers (2018).

⁶⁶. Voir par exemple Casalis & Colé (2018) et Peereman & Sprenger-Charolles (2018).

⁶⁷. Voir pour une présentation d'outils numérisés dans ce domaine, Lassault & Ziegler (2018)

4.5. Quelques limites de la présente étude

Le pouvoir explicatif des variables qui ont un impact tant sur le niveau de vocabulaire en français L2 que sur la prélecture ou la lecture, est plutôt faible. Par ailleurs, certains facteurs ont été intégrés dans les modèles, sans que leur apport ne fasse de différence significative dans l'explication de la variabilité des niveaux des élèves, tant en vocabulaire oral en français qu'en prélecture ou en lecture. C'est notamment le cas pour leur environnement linguistique. De nombreuses questions ont été posées aux élèves et à leurs parents pour essayer de caractériser cet environnement : langue(s) maternelle(s) parlée(s) à la maison, avec les amis, etc. Seule a pu être prise en compte l'opposition environnement familial monolingue versus plurilingue en raison d'un trop faible effectif d'élèves parlant seulement pular ou seereer dans leur vie quotidienne. Cette variable s'est avérée ne pas avoir d'impact sur les compétences évaluées. Il aurait sans doute été intéressant de disposer d'un échantillon plus grand d'élèves, pour pouvoir examiner d'éventuels liens entre les spécificités des LI des enfants et leur niveau de vocabulaire oral en français.

Enfin, comme déjà signalé (dans la section 4.4.1.3), il n'y a pas eu d'évaluation du niveau de vocabulaire oral dans la LN des élèves, ce qui est regrettable. En effet, cette évaluation aurait pu permettre de mieux répondre à une des questions cruciales pour les politiques éducatives : celle du moment le plus propice pour introduire l'apprentissage de la lecture dans une LN.

4.6. Pour conclure

Les résultats de la présente étude permettent d'avancer que, en début de CE1, la compréhension du langage écrit en français devrait être surtout abordée à l'oral, à partir de textes lus par l'enseignant, en particulier des récits parce que les textes de ce type permettent, quand ils sont lus à plusieurs reprises, de construire une mémoire collective et, de ce fait, une culture collective⁶⁸. Ces relectures permettent également de familiariser les élèves avec les spécificités du français écrit. Bien entendu, la compréhension du français doit être également travaillée à l'oral à partir de différentes situations de communication, incluant des échanges dans la classe à propos des textes qui ont été lus par l'enseignant.

La présente étude a permis de valider de nombreux résultats conformes aux données de la littérature internationale. C'est le cas pour ceux qui signalent l'incidence positive du milieu socioculturel sur le niveau de vocabulaire en français. Plus intéressant est l'observation des effets positifs, sur ce niveau, d'activités pratiquées au sein de la famille (lire, parler et apprendre des mots en français à l'enfant), activités qu'il faut développer. Il en est de même pour les données concernant la participation des élèves au programme « Lecture pour tous » (LPT), qui ont donc bénéficié d'heures d'enseignement de lecture en langue nationale. Ces élèves ont en effet un niveau de lecture et de vocabulaire en français supérieur d'environ la moitié d'un écart-type à ceux qui n'ont pas bénéficié de ce dispositif.

En dehors du programme Lecture Pour Tous, ce qui ressort des évaluations des effets des pratiques des enseignants sur les résultats des élèves, pour le vocabulaire comme pour la lecture, est décevant. Ces résultats signalent qu'il est nécessaire d'améliorer la formation (initiale et continue) des maîtres afin de leur permettre d'intensifier, et de diversifier, les

⁶⁸. Bruner (2002).

activités scolaires autour de la compréhension du français oral (du mot au texte), avec un cadre de référence approprié. Il en est de même pour ce qui concerne l'apprentissage du code écrit, dans les LN des enfants, et en français. Les activités dans ce domaine doivent être systématiques, intensives et précoces, l'entrée dans l'écrit se faisant d'abord dans les LN des élèves, parce qu'ils maîtrisent mieux cette langue à l'oral que le français. De plus, la plupart des LN du Sénégal ont une orthographe plus régulière que celle du français, ce qui va faciliter l'apprentissage des correspondances graphème-phonème et le transfert à une orthographe moins régulière : celle du français. C'est bien ce que suggère les résultats des élèves ayant participé au programme Lecture Pour Tous⁶⁹.

Pour dépasser certaines des limites du présent projet, il serait bien de pouvoir effectuer de nouvelles investigations. D'une part, il faudrait s'assurer que, en CE2, le niveau des élèves en compréhension orale du français (mots et groupes de mots) est suffisant pour leur permettre de comprendre ce qu'ils lisent dans cette langue. D'autre part, il faudrait pouvoir vérifier les relations, dans les LN des enfants, entre le niveau de vocabulaire oral et le niveau de lecture.

⁶⁹. Les interventions précoces de type « Lecture Pour Tous » sont d'autant plus importantes que les différences initiales dues au milieu socioculturel s'accroissent dans le temps, cf. données de PIRLS, voir aussi Baudelot & Establet (2009) ; Broccolichi & Sinthon (2011).

5. BIBLIOGRAPHIE

Abadzi, H. (2007). *Absenteeism and Beyond: Instructional Time Loss and Consequences*. Washington DC: Banque Mondiale.

Anderson R. C., Wilson P. T. & Fielding L. G. (1988). Growth in reading and how children spend their time outside of school. *Reading Research Quarterly*, 23, 285–303. doi:10.1598/rrq.23.3.2

Arnauld A. & Lancelot C. (1660). *Grammaire générale et raisonnée de Port-Royal*. Paris : Bossange et Masson (Seconde édition, 1810).

August D. & Shanahan T. (2006). *Developing literacy in second-language learners*. Washington, DC: Lawrence Erlbaum Associates and the Center for Applied Linguistics.

Baudelot C. & Establet R. (2009), *L'élitisme républicain. L'école française à l'épreuve des comparaisons internationales*, Paris : Seuil

Broccolichi S. & Sinthon R. (2011). Comment s'articulent les inégalités d'acquisition scolaire et d'orientation ? Relations ignorées et rectifications tardives. *Revue Française de Pédagogie*, 175, 15-38.

Bernard J.-M., Tiyaab K., Vianou K. (2004). *Profils enseignants et qualité de l'éducation primaire en Afrique subsaharienne francophone : Bilan et perspectives de dix années de recherche du PASEC*. Dakar : PASEC/CONFEMEN.

Braga M., Checchi D., Garrouste C. et al. (2019). *Selecting or Rewarding Teachers? International Evidence from Primary Schools*, IZA DP 12200 <https://www.iza.org/publications/dp/12200/selecting-or-rewarding-teachers-international-evidence-from-primary-schools>

Bressoux P. (2001). Réflexions sur l'effet-maître et l'étude des pratiques enseignantes. *Les dossiers des sciences de l'éducation*, 5, pp. 35-52.

Bruner J. (2002). *Pourquoi nous racontons-nous des histoires*. Paris : Retz.

Caillies S. (2009). Descriptions de 300 expressions idiomatiques, *Année Psychologique*, 109(3), 463-508.

Casalis S. & Colé P. (2018). Le morphème, une unité de traitement dans l'acquisition de la littéracie, *Langue française*, 199, 69-81.

Castles A., Rastle K. & Cain K. (2018). Ending the Reading Wars: Reading acquisition from novice to expert. *Psycholog Science in Public Interest*, 19(1), 5-51 [<http://journals.sagepub.com/doi/abs/10.1177/1529100618772271>].

Charolles M. (1995). Cohésion, Cohérence et pertinence du discours. *Travaux de Linguistique*, 29, 125-151.

Clifton C. Jr., Ferreira F., Henderson J. M. et al. (2016). Eye movements in reading and information processing: Keith Rayner's 40 years legacy. *Journal of Memory and Language*, 86, 1-19

Currie N. K. & Cain K. (2015). Children's inference generation: The role of vocabulary and working memory. *Journal of Experimental Child Psychology*, 137, 57–75

Deacon S. H., Desrochers A. & Levesque K. (2017). *Learning to Read French*. in L. Verhoeven & C. Perfetti (eds.), *Learning to Read across Languages and Writing Systems* (pp.243-269), Cambridge University Press.

- Dehaene S. (Dir) (2011). *Apprendre à lire. Des sciences cognitives à la salle de classe*. Paris : Odile Jacob.
- Dehaene S. (Dir) (2019). La science au service de l'école (Chapitre 2: Evaluer pour mieux aider, pp. 85-158), Odile Jacob.
- Delattre P. (1965). *Comparing the phonetic features of English, French, German and Spanish*. Heidelberg: Jumiuss Gross Verlag.
- Deltour J.J. & Hupkens D. (1980). *Test de vocabulaire actif et passif pour enfants de 5 à 8 ans (TVAP 5-8)*. Braine-le-Château: Editions de l'Application des Techniques Modernes.
- Dunn L. M., Thériault-Whalen C. M. & Dunn L. M. (1993). *Echelle de vocabulaire en image du Peabody*. Psycan, Toronto.
- Early Literacy Panel (2008). *Developing early literacy: A Scientific Synthesis of Early Literacy Development and Implications for Intervention* <http://lincs.ed.gov/publications/pdf/NELPReport09.pdf>
- Fayol M. (2017). *L'acquisition de l'écrit*. Paris: Presses Universitaires de France.
- Gentaz E., Sprenger-Charolles L., Theurel A. et al. (2013). Reading comprehension in a large cohort of French first graders from low socio-economic status families: A 7-month longitudinal study. *Plos ONE*, 8(11): e78608.2.
- Gentaz E., Sprenger-Charolles L. & Theurel A. (2015). Differences in the predictors of reading comprehension in first graders from low socio-economic status families with either good or poor decoding skills. *PLoS ONE*. 10(3): e0119581.
- Gernsbacher M., Varner K. & Faust M. (1990). Investigating differences in general comprehension skills. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 16, 430-445.
- Hoff E. (2006). How social contexts support and shape language development. *Developmental Review* 26: 55–88.
- Hoover W. A. & Gough P. B (1990). The simple view of reading. *Reading and Writing: An Interdisciplinary Journal*, 2, 127-160.
- Kolinsky R., Morais J., Cohen L. et al. (2018). Les bases neurales de l'apprentissage de la lecture, *Langue française*, 199, 17-33.
- Lassault J. & Ziegler J.C. (2018). Les outils numériques d'aide à l'apprentissage de la lecture. *Langue française*, 199, 111-121
- Lecocq P. (1996a). *ECosSe (Epreuve de Compréhension Syntaxico-Sémantique)*. Lille: Presses du Septentrion.
- Lecocq P., Casalis S., Leuwens C. et al. (1996b). *Apprentissage de la lecture et compréhension d'énoncés*. Lille: Presses universitaires du Septentrion.
- Lété B, Sprenger-Charolles L., & Colé P. (2004). MANULEX: A lexical database from French readers. *Behavioral Research Methods, Instruments and Computers*, 36(1), 156-166. <http://www.manulex.org/>.

- Michaelowa, K., Wechtler, A. (2006). *Coût-efficacité des intrants de l'enseignement primaire : ce que nous apprend la documentation et des récentes enquêtes sur les étudiants en Afrique subsaharienne*. Papier présenté à la Biennale de l'ADEA, Libreville, Gabon.
- Mingat A., Suchaut B. (2000). *Les systèmes éducatifs africains : une analyse économique comparative*. Bruxelles : De Boeck.
- Murphy V. A. (2018). *Literacy development in linguistically diverse pupils*. In D. Miller et al. (Eds.), *Bilingual cognition and language: The state of the science across its subfields*. *Studies in Bilingualism* (Vol. 54, pp. 155–182). Amsterdam, The Netherlands: John Benjamins. doi:10.175/sibil.54.08mur
- Nancarrow O. (2004). *L'Afrique*. In B. Comrie, S. Matthews & M. Polinsky (Eds). *Atlas des langues : L'origine et le développement des langues dans le monde* (pp.72-89). Paris, Edition Acropole.
- National curriculum in England (2013). *English programs of study: key stages 1 and 2*. National archives https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/335186/PRIMARY_national_curriculum_-_English_220714.pdf
- Noble K.G. & McCandliss B.D. (2005). Reading development and impairment: behavioral, social, and neurobiological factors. *Journal of Developmental and Behavioral Pediatrics* 26: 370–378.
- Peereman R. & Sprenger-Charolles L. (2018). Manulex-MorphO, une base de données sur l'orthographe du français intégrant les morpho-phonogrammes. *Langue Française*, 199.
- Perfetti C. A. (2007). Reading Ability: Lexical Quality to Comprehension. *Scientific Studies of Reading* 11(4):357-383.
- Perfetti C. A. & Stafura J. (2014). Word knowledge in a theory of reading comprehension. *Scientific Studies of Reading*, 18, 22–37.
- Piquard-Kipffer A. & Sprenger-Charolles L. (2013). Early predictors of future reading skills: A follow-up of French-speaking children from the beginning of kindergarten to the end of the second grade (age 5 to 8). *Topics in Cognitive Psychology – L'Année Psychologique*, 113, 491-522.
- PIRLS [Progress in International Reading Literacy Study] (2017). *PIRLS 2016*. IEA [International Association for the Evaluation of Educational Achievement]. <http://www.iea.nl/pirls>.
- Quinn J. M., Wagner R. K., Petscher Y. et al. (2015). Developmental relations between vocabulary knowledge and reading comprehension: A latent change score modeling study. *Child Development*, 86, 159–175.
- Riou J. & Fontanieu V. (2016). Influence de la planification de l'étude du code alphabétique sur les performances des élèves en décodage au cours préparatoire. *Revue française de pédagogie*. 196, 49–66. <https://doi.org/10.4000/rfp.5073>
- Rose J. (2006). *Independent review of the teaching of early reading final report*. U.K. Department for Education and Skills.
- RTI International (2016). *Manuel pour l'évaluation des compétences fondamentales en lecture* (Adaptation au français de EGRA [Early Grade Reading Assessment], 2nd édition). Washington, DC, Agence Américaine pour le développement international

https://www.globalreadingnetwork.net/sites/default/files/resource_files/EGRA_Toolkit_2nd_Ed_French-Final.pdf

Share D. L. (1995). Phonological recoding and self-teaching: Sine qua non of reading acquisition. *Cognition*, 55, 151-218.

Sprenger-Charolles L. (2008). *Rapport sur l'étude pilote EGRA Sénégal*, Banque Mondiale

Sprenger-Charolles L. & Desrochers A. (Eds.) (2018). Lire-Ecrire : des savoirs scientifiques aux savoirs pratiques. *Langue Française*, 199, 3.

Sprenger-Charolles L., Siegel L. S., Béchenec D. et al. (2003). Development of phonological and orthographic processing in reading aloud, in silent reading and in spelling: A four-year longitudinal study. *Journal of Experimental Child Psychology*, 84, 194-217.

Torgesen J. K. (2009). The Response to Intervention Instructional Model: Some Outcomes from a Large-scale Implementation in Reading First Schools. *Child Development Perspectives*, 3(1), 38-40.

Treiman R., Tincoff R. & Richmond-Welty E. D. (1996). Letter names help children to connect print and speech. *Developmental Psychology*, 32(3), 505–514. <https://doi.org/10.1037/0012-1649.32.3.505>

Treiman R. & Tincoff R. (1997). The fragility of the alphabetic principle: Children's knowledge of letters names can cause them to spell syllabically rather alphabetically. *Journal of Experimental Child Psychology*, 64, 425–451.

Tunmer W.E. & Chapman J.W. (2012). Does set for variability mediate the influence of vocabulary knowledge on the development of word recognition skills? *Scientific Studies of Reading*, 16(2), 122-140.

Ziegler J. C. (2018). Différences inter-linguistiques dans l'apprentissage de la lecture, *Langue française*, 199, 35-49.

6. ANNEXES

6.1. Protocole d'évaluation

DONNER LA CONSIGNE EN LANGUE NATIONALE SAUF LE MOT QUE L'ITEM CHERCHE A EVALUER, QUI DOIT ABSOLUMENT RESTER EN FRANÇAIS.

Consigne de passation	Durée	Barème de notation
-----------------------	-------	--------------------

I. Vocabulaire oral passif : la désignation d'images (mots isolés)

Planche I

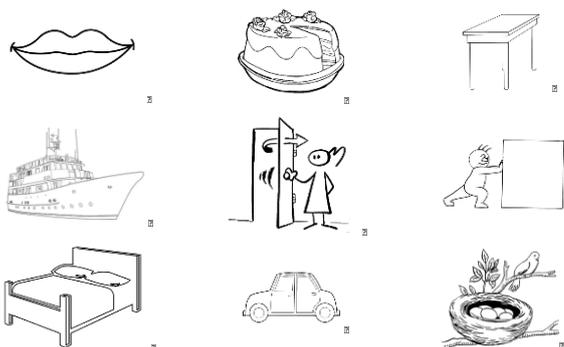


Planche I

Consigne générale / Test

Regarde bien ces images. Montre-moi la bouche.

Montre-moi celle qui correspond à l'action de pousser.

Item I.1.1

Regarde bien ces images. Montre-moi l' <u>auto</u> .	15 secondes	0	1	9
		Mauvaise réponse	Bonne réponse	Pas de réponse

Item I.1.2

Montre-moi le <u>gâteau</u> .	15 secondes	0	1	9
		Mauvaise réponse	Bonne réponse	Pas de réponse

Item I.1.3

Montre-moi celle qui correspond à l'action d' <u>ouvrir</u> .	15 secondes	0	1	9
		Mauvaise réponse	Bonne réponse	Pas de réponse

Item I.1.4

Montre-moi le <u>lit</u> .	15 secondes	0	1	9
		Mauvaise réponse	Bonne réponse	Pas de réponse

Item I.1.5

Montre-moi la <u>table</u> .	15 secondes	0	1	9
		Mauvaise réponse	Bonne réponse	Pas de réponse

Item 1.1.6 : Arrêt si l'enfant n'a donné qu'1 réponse correcte sur les 5 items => passer à la section 2	0	1
	Non concerné	Concerné

Planche 2

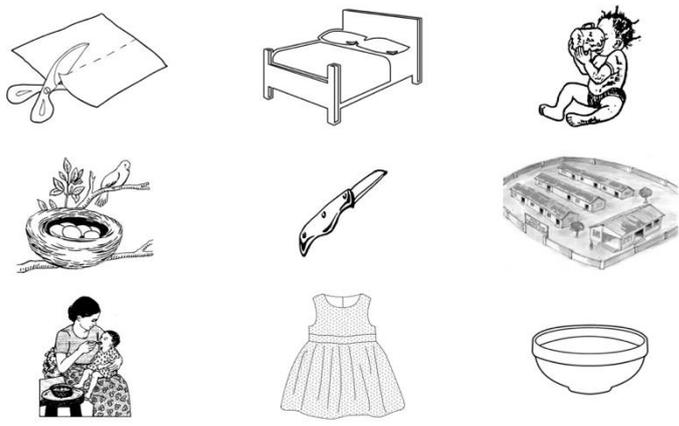


Planche 2

Item 1.2.1				
Regarde bien ces images. Montre-moi le <u>bol</u> .	15 secondes	0	1	9
		Mauvaise réponse	Bonne réponse	Pas de réponse
Item 1.2.2				
Montre-moi celle qui correspond à l'action de <u>boire</u> .	15 secondes	0	1	9
		Mauvaise réponse	Bonne réponse	Pas de réponse
Item 1.2.3				
Montre-moi l' <u>école</u> .	15 secondes	0	1	9
		Mauvaise réponse	Bonne réponse	Pas de réponse
Item 1.2.4				
Regarde bien ces images. Montre-moi le <u>nid</u> .	15 secondes	0	1	9
		Mauvaise réponse	Bonne réponse	Pas de réponse
Item 1.2.5				
Regarde bien ces images. Montre-moi la <u>robe</u> .	15 secondes	0	1	9
		Mauvaise réponse	Bonne réponse	Pas de réponse

Item 1.2.6 : Arrêt si l'enfant n'a donné qu'1 réponse correcte sur les 5 items => passer à la section 2	0	1
	Non concerné	Concerné

Planche 3

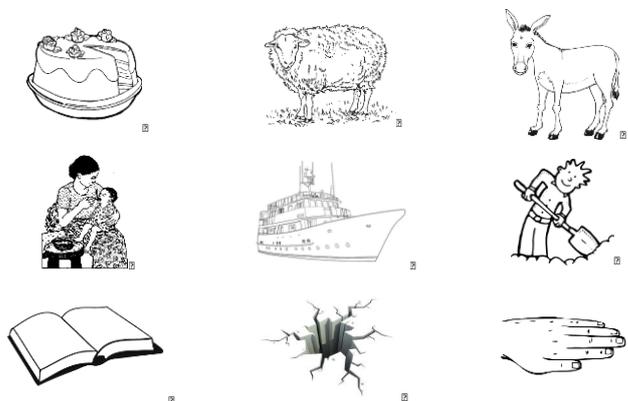


Planche 3

Item 1.3.1				
Regarde bien ces images. Montre-moi celle qui correspond à l'action de manger .	15 secondes	0	1	9
		Mauvaise réponse	Bonne réponse	Pas de réponse
Item 1.3.2				
Montre-moi le bateau .	15 secondes	0	1	9
		Mauvaise réponse	Bonne réponse	Pas de réponse
Item 1.3.3				
Montre-moi le mouton .	15 secondes	0	1	9
		Mauvaise réponse	Bonne réponse	Pas de réponse
Item 1.3.4				
Montre-moi le livre .	15 secondes	0	1	9
		Mauvaise réponse	Bonne réponse	Pas de réponse
Item 1.3.5				
Montre-moi celle qui correspond à l'action de creuser .	15 secondes	0	1	9
		Mauvaise réponse	Bonne réponse	Pas de réponse
Item 1.3.6				
Montre-moi la main .	15 secondes	0	1	9
		Mauvaise réponse	Bonne réponse	Pas de réponse

Item 1.3.7 : Arrêt si l'enfant n'a donné qu'1 réponse correcte sur les 6 items => passer à la section 2	0	1
	Non concerné	Concerné

Planche 4

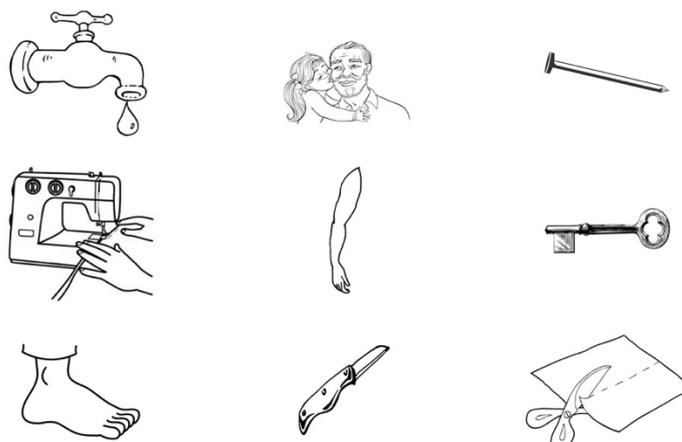


Planche 4

Item I.4.1				
Regarde bien ces images. Montre-moi le robinet .	15 secondes	0	1	9
		Mauvaise réponse	Bonne réponse	Pas de réponse
Item I.4.2				
Montre-moi le bras .	15 secondes	0	1	9
		Mauvaise réponse	Bonne réponse	Pas de réponse
Item I.4.3				
Montre-moi celle qui correspond à l'action de coudre .	15 secondes	0	1	9
		Mauvaise réponse	Bonne réponse	Pas de réponse
Item I.4.4				
Montre-moi la clé .	15 secondes	0	1	9
		Mauvaise réponse	Bonne réponse	Pas de réponse
Item I.4.5				
Montre-moi le couteau .	15 secondes	0	1	9
		Mauvaise réponse	Bonne réponse	Pas de réponse
Item I.4.6				
Montre-moi celle qui correspond à l'action de couper .	15 secondes	0	1	9
		Mauvaise réponse	Bonne réponse	Pas de réponse

2. Vocabulaire oral actif : mots isolés et en contexte

Items 2.1 : Les jours de la semaine (6 points + 1 item d'arrêt)

Item 2.1.1				
Je vais te poser des questions sur les jours de la semaine. Quel jour vient après le lundi ?	10 secondes	0	1	9
		Mauvaise réponse	Bonne réponse	Pas de réponse
Item 2.1.2				
Quel jour vient après le mardi ?	10 secondes	0	1	9
		Mauvaise réponse	Bonne réponse	Pas de réponse

Item 2.1.3 : Arrêt si l'enfant n'a donné aucune réponse => Passer à la section 2.2	0	1
	Non concerné	Concerné

Item 2.1.4				
Quel jour vient après le mercredi ?	10 secondes	0	1	9
		Mauvaise réponse	Bonne réponse	Pas de réponse
Item 2.1.5				
Quel jour vient après le jeudi ?	10 secondes	0	1	9
		Mauvaise réponse	Bonne réponse	Pas de réponse
Item 2.1.6				
Quel jour vient après le vendredi ?	10 secondes	0	1	9
		Mauvaise réponse	Bonne réponse	Pas de réponse
Item 2.1.7				
Quel jour vient après le samedi ?	10 secondes	0	1	9
		Mauvaise réponse	Bonne réponse	Pas de réponse

Items 2.2 : Les parties du corps (7 points + 1 item d'arrêt)

Item 2.2.1					
Montre-moi ton nez .	10 secondes	0	1		9
		Mauvaise réponse	Bonne réponse		Pas de réponse
Item 2.2.2					
Montre-moi ta bouche .	10 secondes	0	1		9
		Mauvaise réponse	Bonne réponse		Pas de réponse
Item 2.2.3					
Montre-moi ta langue .	10 secondes	0	1		9
		Mauvaise réponse	Bonne réponse		Pas de réponse

Item 2.2.5					
Montre-moi ton <u>cou</u> .	10 secondes	0	1		9
		Mauvaise réponse	Bonne réponse		Pas de réponse
Item 2.2.6					
Montre-moi ta <u>main droite</u> .	15 secondes	0	1	2	9
		Mauvaise réponse	Montre sa main gauche	Montre sa main droite	Pas de réponse
Item 2.2.7					
Montre-moi ton <u>coude gauche</u> .	15 secondes	0	1	2	9
		Mauvaise réponse	Montre son coude droit.	Montre son coude gauche.	Pas de réponse
Item 2.2.8					
Montre-moi ton <u>épaule droite</u> .	15 secondes	0	1	2	9
		Mauvaise réponse	Montre son épaule gauche.	Montre son épaule droite.	Pas de réponse

Items 2.3 : Les verbes d'action (8 points)

Item 2.3.1					
Je vais te demander de faire des actions. Je te demande de <u>rire</u> .	10 secondes	0	1	9	
		Mauvaise réponse	Bonne réponse	Pas de réponse	
Item 2.3.2					
Je te demande de <u>tousser</u> .	10 secondes	0	1	9	
		Mauvaise réponse	Bonne réponse	Pas de réponse	

Item 2.3.3 : Arrêt si l'enfant n'a donné aucune réponse		0	1
Fin cette partie du test : passer à la partie 3 (Lecture)		Non concerné	Concerné

Item 2.3.4					
Je te demande de tracer une <u>ligne</u> avec ton doigt.	10 secondes	0	1	9	
		Mauvaise réponse	Bonne réponse	Pas de réponse	
Item 2.3.5					
Je te demande de tracer une <u>croix</u> avec ton doigt.	10 secondes	0	1	9	
		Mauvaise réponse	Bonne réponse	Pas de réponse	
Item 2.3.6					
Je te demande de <u>respirer</u> fort.	10 secondes	0	1	9	
		Mauvaise réponse	Bonne réponse	Pas de réponse	
Item 2.3.7					
Je te demande de faire une <u>grimace</u> .	10 secondes	0	1	9	
		Mauvaise réponse	Bonne réponse	Pas de réponse	

Items 2.4 : Espace - Temps (8 points)

Item 2.4.1					
Tiens, je te donne un objet. Mets cet objet devant toi .	15 secondes	0	1		9
		Agis mais ne mets pas le caillou devant lui	Mets le caillou devant lui		Pas de réponse
Item 2.4.2					
Maintenant, mets cet objet devant moi .	15 secondes	0	1		9
		Agis mais ne mets pas le caillou derrière moi	Mets le caillou devant moi		Pas de réponse
Item 2.4.3					
Maintenant, mets cet objet derrière toi .	15 secondes	0	1		9
		Agis mais ne mets pas le caillou derrière lui	Mets le caillou derrière lui		Pas de réponse
Item 2.4.4					
Maintenant, mets cet objet près de moi .	15 secondes	0	1		9
		Agis mais ne mets pas le caillou près de moi	Mets le caillou près de moi		Pas de réponse
Item 2.4.5					
Maintenant, mets cet objet loin de moi .	15 secondes	0	1		9
		Agis mais ne mets pas le caillou loin de moi	Mets le caillou loin de moi		Pas de réponse
Item 2.4.6					
Maintenant, mets cet objet dans ta main droite .	15 secondes	0	1	2	9
		Agis mais ne mets pas le caillou dans une de ses mains	Mets le caillou dans sa main gauche	Mets le caillou dans sa main droite	Pas de réponse

3. Épreuve de prélecture et lecture

Seulement pour les élèves de CI et de CP

Je vais te demander de lire des petits bouts de mots. Je te montre un exemple.
Montrer la lettre « o » et dire « o ». A ton tour maintenant.
Si l'élève ne répond pas après 3 secondes, passer à l'item suivant.

Item	« Bouts » de mots à lire	Mauvaise réponse	Se trompe mais se corrige	Bonne réponse	Pas de réponse		
3.1.1	a	0	1	2	9		
3.1.2	i	0	1	2	9		
3.1.3	é	0	1	2	9		
3.1.4	o	0	1	2	9		
3.1.5	S'arrêter si l'enfant n'a lu aucune des lettres						99
3.1.6	sa	0	1	2	9		
3.1.7	la	0	1	2	9		
3.1.8	dé	0	1	2	9		
3.1.9	ti	0	1	2	9		

Seulement pour les élèves de CE1

Je vais te demander de lire des petits bouts de mots. Je te montre un exemple.

Montrer la lettre « o » et dire « o ». A ton tour maintenant.

L'épreuve se déroule sur 1 minute et 30 secondes.

Si l'élève ne répond pas après 3 secondes, passer à l'item suivant.

Item	Mot à lire	Mauvaise réponse	Se trompe mais se corrige	Bonne réponse	Pas de réponse	Arrêt au bout d'1 minute 30
3.2.1	sa	0	1	2	9	99
3.2.2	la	0	1	2	9	99
3.2.3	dé	0	1	2	9	99
3.2.4	ti	0	1	2	9	99
3.2.5	al	0	1	2	9	99
3.2.6	fo	0	1	2	9	99
3.2.7	tifo	0	1	2	9	99
3.2.8	lato	0	1	2	9	99
3.2.9	tafi	0	1	2	9	99
3.2.10	sabol	0	1	2	9	99
3.2.11	daba	0	1	2	9	99
3.2.12	oto	0	1	2	9	99
3.2.13	bato	0	1	2	9	99
3.2.14	foto	0	1	2	9	99
3.2.15	mati	0	1	2	9	99
3.2.16	alima	0	1	2	9	99
3.2.17	bal	0	1	2	9	99
3.2.18	ami	0	1	2	9	99
3.2.19	midi	0	1	2	9	99
3.2.20	kilo	0	1	2	9	99
3.2.21	moto	0	1	2	9	99
3.2.22	mali	0	1	2	9	99

6.2. Questionnaire Élève

1. Identification de l'école élémentaire et de l'élève

1. Nom de l'école :	2. Code :
---------------------	-----------

3. IA	4. IEF	5. Commune

6. Prénom	7. Nom	8. Classe	9. Code

2. Questions sur l'environnement linguistique de l'enfant

10. Quelle langue parles-tu à la maison ? <i>Plusieurs réponses possibles.</i>				
Pulaar	Seereer	Wolof	Français	Autre :

11. Quelle langue parles-tu avec tes amis ? <i>Plusieurs réponses possibles.</i>				
Pulaar	Seereer	Wolof	Français	Autre :

12. Quelle langue parles-tu avec ton enseignant ? <i>Plusieurs réponses possibles.</i>				
Pulaar	Seereer	Wolof	Français	Autre :

13. Regardes-tu la télévision ?	Oui	Non (passez à la question ...)
---------------------------------	-----	--------------------------------

14. Si oui, dans quelle(s) langue(s) sont les émissions que tu regardes ? <i>Plusieurs réponses possibles.</i>				
Pulaar	Seereer	Wolof	Français	Autre :

15. Y'a-t-il au moins une personne à la maison qui sait lire ?	Oui	Non	Je ne sais pas
--	-----	-----	----------------

16. Est-ce que tu regardes ou lis des livres avec cette personne ?	Oui	Non
--	-----	-----

17. Est-ce qu'il y a des livres que tu peux regarder/lire à la maison ?	Oui	Non
---	-----	-----

18. Si oui, passes-tu du temps à regarder ou lire ces livres ?				
Non, pas du tout	Une fois par mois	Quelques fois par mois	Quelques fois par semaine	Tous les jours

19. Y'a-t-il une personne à la maison qui parle français ?	Oui	Non	Je ne sais pas
--	-----	-----	----------------

20. Si oui, parles-tu français avec elle ?				
Non, pas du tout	Une fois par mois	Quelques fois par mois	Quelques fois par semaine	Tous les jours

21. Si oui, est-ce que cette personne t'apprend des choses en français (des mots, des expressions...) ?				
Non, pas du tout	Une fois par mois	Quelques fois par mois	Quelques fois par semaine	Tous les jours

6.3. Questionnaire Parent d'élève

1) Comment s'appelle votre enfant ?

2) Quel est le sexe de votre enfant ?

Garçon	Fille
--------	-------

3) Quelle est sa date de naissance ?

Jour :	Mois :	Année :
--------------	--------------	---------------

4) Dans quelle classe est actuellement votre enfant ?

CI (passer à la question 6)	CP	CEI
-----------------------------	----	-----

5) Si votre enfant est scolarisé en CP ou en CEI, était-il scolarisé dans la même école l'année dernière ?

Oui	Non
-----	-----

6) Avant d'aller à l'école élémentaire, votre enfant a-t-il fréquenté le préscolaire ?

Oui	Non (passer à la question 8)
-----	------------------------------

7) Si oui, combien d'années l'a-t-il fréquenté ? Et quel type de préscolaire ?

Moins d'1 an	1 an	2 ans	3 ans ou plus
--------------	------	-------	---------------

Communaire	Case des Tout Petits	Classe maternelle	École maternelle			Garderie	Daara préscolaire
			Publique	Privée	Privée Franco-Arabe		

8) Votre enfant vit-il avec ses deux parents ?

Oui	Non, avec sa mère	Non, avec son père	Non, aucun des deux
-----	-------------------	--------------------	---------------------

9) Votre enfant a-t-il des frères et sœurs, avec lesquels il vit ? Plusieurs réponses possibles.

Oui, plus âgés	Oui, plus jeunes	Non
----------------	------------------	-----

10) Savez-vous lire ? Et son autre parent ?

Père ou tuteur		Mère ou tutrice	
Oui	Non	Oui	Non

11) Quel est le plus haut niveau d'instruction que vous avez atteint ? Et celui de son autre parent ?

Père ou tuteur	Mère ou tutrice
Non scolarisé	Non scolarisé
Alphabétisation	Alphabétisation
Élémentaire	Élémentaire
Collège	Collège
Lycée	Lycée
Supérieur	Supérieur

12) Quelle est votre profession ? Et celle de son autre parent ?

Père ou tuteur	Mère ou tutrice
Agriculteur	Agriculteur
Vendeur	Vendeur
Commerçant, Artisan	Commerçant, Artisan
Chauffeur	Chauffeur
Cadre	Cadre
Employé, ouvrier	Employé, ouvrier
Retraité	Retraité
Sans emploi	Sans emploi
A l'étranger	A l'étranger
Autre	Autre

13) Quelle est votre langue maternelle ? Et celle de son autre parent ?

Père ou tuteur				Mère ou tutrice			
Wolof	Pulaar	Seereer	Autre	Wolof	Pulaar	Seereer	Autre

14) Quelle(s) langue(s) parlez-vous à la maison ? Plusieurs réponses possibles.

Père ou tuteur					Mère ou tutrice				
Wolof	Pulaar	Seereer	Français	Autre	Wolof	Pulaar	Seereer	Français	Autre

15) S'agissant du français :

	Père ou tuteur			Mère ou tutrice		
Comprenez-vous le français ?	Oui	Un peu	Non	Oui	Un peu	Non
Parlez-vous le français ?	Oui	Un peu	Non	Oui	Un peu	Non
Savez-vous lire en français ?	Oui	Un peu	Non	Oui	Un peu	Non

16) A quelle fréquence parlez-vous français à votre enfant ? Et son autre parent ?

Père ou tuteur	Mère ou tutrice
Jamais	Jamais
Quelques fois par mois	Quelques fois par mois
Quelques fois par semaine	Quelques fois par semaine
Tous les jours ou presque	Tous les jours ou presque

6.3. Questionnaire Enseignant

Dans chaque école, deux enseignants doivent être interrogés :

- . L'enseignant qui tenait la classe de CI en 2017-2018
- . L'enseignant qui tenait la classe de CP en 2017-2018

1. Identification de l'école élémentaire

Nom de l'école :	Code :
------------------	--------

IA	IEF	Commune

2. Identification et caractéristiques de l'enseignant

Niveau enseigné en 2017-2018	CI	CP
------------------------------	----	----

Sexe	Homme	Femme
------	-------	-------

Langue maternelle	Pulaar	Seereer	Wolof	Autre :
-------------------	--------	---------	-------	---------------

Statut	Fonctionnaire	Contractuel	Volontaire
--------	---------------	-------------	------------

Plus haut niveau académique atteint								
< 3 ^{ème}	3 ^{ème}	2 nd e	1 ^{ère}	Terminale	Bac +1	Bac +2	Bac +3	Bac +4 ou +

Plus haut diplôme obtenu	BFEM/DFEM	Baccalauréat	Licence	Maîtrise	Autre	Aucun
--------------------------	-----------	--------------	---------	----------	-------	-------

Diplôme professionnel	Aucun	CEAP	CAP	Autre
-----------------------	-------	------	-----	-------

Année de début de carrière	
----------------------------	--

Nombre de jours de formation continue ces dernières années (tous types et formes confondus)							
2010-11	2011-12	2012-13	2013-14	2014-15	2015-16	2016-2017	2017-18

3. Perceptions de l'aisance en français

Dans quelle mesure pensez-vous être à l'aise pour lire en français ?				
Pas du tout à l'aise	Peu à l'aise	Moyennement à l'aise	Plutôt à l'aise	Très à l'aise

Dans quelle mesure pensez-vous être à l'aise pour écrire en français ?				
Pas du tout à l'aise	Peu à l'aise	Moyennement à l'aise	Plutôt à l'aise	Très à l'aise

Dans quelle mesure pensez-vous être à l'aise pour parler en français ?				
Pas du tout à l'aise	Peu à l'aise	Moyennement à l'aise	Plutôt à l'aise	Très à l'aise

Dans quelle mesure pensez-vous être à l'aise pour comprendre en français ?				
Pas du tout à l'aise	Peu à l'aise	Moyennement à l'aise	Plutôt à l'aise	Très à l'aise

Dans quelle mesure pensez-vous être à l'aise pour enseigner en français ?				
Pas du tout à l'aise	Peu à l'aise	Moyennement à l'aise	Plutôt à l'aise	Très à l'aise

4. Questions sur la classe dans laquelle il/elle enseignait en 2017-2018

4.1. Caractéristiques de la classe

Nombre d'élèves dans la classe	
--------------------------------	--

Votre classe faisait-elle partie d'un programme particulier ?	Oui	Non
---	-----	-----

Si oui, quel est ce programme ?					
LPT	ELAN	ARED / Emil	IFADEM	ADLAS	Autre :

	Manuels disponibles				Si oui, sont-ils utilisés ?
	Oui, 1 par élève	Oui, mais 1 livre pour 2 élèves	Oui, mais 1 livre pour 3 élèves ou plus	Non	
Manuel de lecture en français					
Livre d'exercices					

À quelle date ont effectivement commencé les enseignements-apprentissages en 2017-18 ?	.. / .. /
--	----------------

À quelle date les enseignements-apprentissages se sont-ils terminés en 2017-18 ?	.. / .. /
--	----------------

Au cours de l'année 2017-18 et en dehors des vacances scolaires, les activités d'enseignements-apprentissages ont-elles connu une interruption ?				
Jamais	< 15 jours	15 jours à 1 mois	1 à 2 mois	> 2 mois

4.2. Situation linguistique dans la classe en 2017-2018

Hormis le français, parliez-vous une autre langue pendant votre enseignement, dans la classe ?	Oui	Non
--	-----	-----

Si oui, quelle(s) langue(s) parliez-vous ? <i>Plusieurs réponses possibles.</i>			
Pulaar	Seereer	Wolof	Autre :

Si oui et que plusieurs langues sont citées : pourriez-vous indiquer la langue que vous utilisez le plus ?			
Pulaar	Seereer	Wolof	Autre :

En dehors de la classe, parliez-vous une autre langue que le français avec vos élèves (pour annoncer la récréation, reprendre un élève sur son comportement, etc.) ?	Oui	Non
--	-----	-----

Si oui, quelle(s) langue(s) parliez-vous ? <i>Plusieurs réponses possibles.</i>			
Pulaar	Seereer	Wolof	Autre :

En prenant en compte l'ensemble du temps passé avec les élèves (dans la classe, pendant la récréation, etc.), diriez-vous que les interactions en <u>français</u> avec vos élèves étaient :				
Très peu fréquentes : 20 % du temps ou -	Peu fréquentes : 20 à 40 %	Moyennement fréquentes : 40 à 60 %	Plutôt fréquentes : 60 à 80 %	Très fréquentes : 80 à 100 %

En prenant en compte l'ensemble du temps passé avec les élèves, diriez-vous que les interactions <u>dans une autre langue que le français</u> étaient :				
Très peu fréquentes : 20 % du temps ou -	Peu fréquentes : 20 à 40 %	Moyennement fréquentes : 40 à 60 %	Plutôt fréquentes : 60 à 80 %	Très fréquentes : 80 à 100 %

4.3. Temps alloué à différentes activités d'enseignement dans la classe en 2017-2018

Dans une semaine-type, combien de <u>minutes</u> consacriez-vous à l'enseignement de la langue orale (en français) ?	
--	--

Dans une semaine-type, combien de <u>minutes</u> consacriez-vous au travail autour du texte et du discours en français (lecture et discussions autour d'un récit, récitation, etc.) ?	
---	--

Dans une semaine-type, combien de <u>minutes</u> consacriez-vous à l'enseignement de la phrase et de la grammaire (en français) ?	
---	--

Dans une semaine-type, combien de <u>minutes</u> consacriez-vous à l'enseignement du vocabulaire oral (en français) ?	
---	--

Dans une semaine-type, combien de <u>minutes</u> consacriez-vous l'apprentissage de la lecture et de l'écriture (en français) ?	
---	--

Plus précisément pour la <u>lecture</u> , dans une semaine-type, combien de <u>minutes</u> consacriez-vous :	
A l'enseignement du code	
Au travail de compréhension	

Plus précisément pour la <u>écriture</u> , dans une semaine-type, combien de <u>minutes</u> consacriez-vous :	
A l'enseignement du code	
Au travail graphique	

4.4. Activités d'enseignement dans la classe en 2017-2018 (en français)

A quelle fréquence lisiez-vous à haute voix des récits aux enfants ?			
Jamais	Une fois par mois	Au moins une fois par semaine	Tous les jours

A quelle fréquence demandiez-vous aux enfants ce qu'ils ont compris du récit que vous leur avez lu (qui étaient les personnages, ce qu'ils ont fait, quand, pourquoi, avec quel résultat) ?			
Jamais	Une fois par mois	Au moins une fois par semaine	Tous les jours

A quelle fréquence demandiez-vous à un des enfants de la classe de raconter un événement qui lui est arrivé ?			
Jamais	Une fois par mois	Au moins une fois par semaine	Tous les jours

A quelle fréquence demandiez-vous aux autres enfants de poser des questions sur cet événement ?			
Jamais	Une fois par mois	Au moins une fois par semaine	Tous les jours

A quelle fréquence demandiez-vous aux enfants de donner leur avis sur un sujet donné (la nourriture, la ville et la campagne...) ?			
Jamais	Une fois par mois	Au moins une fois par semaine	Tous les jours

A quelle fréquence demandiez-vous aux enfants de réciter par cœur des textes (récit, poésie ...) ?			
Jamais	Une fois par mois	Au moins une fois par semaine	Tous les jours

A quelle fréquence demandiez-vous aux enfants de remettre des mots dans l'ordre (par exemple, « Pape Pomme Mange » => « Pape mange une pomme ») et d'ajouter le mot qui manque ?			
Jamais	Une fois par mois	Au moins une fois par semaine	Tous les jours

A quelle fréquence travailliez-vous avec les enfants sur les types de phrases (déclarative, interrogative, négative, impérative) ?			
Jamais	Une fois par mois	Au moins une fois par semaine	Tous les jours

A quelle fréquence travailliez-vous avec les enfants sur la conjugaison des verbes au présent ?			
Jamais	Une fois par mois	Au moins une fois par semaine	Tous les jours

A quelle fréquence travailliez-vous avec les enfants sur les moyens d'indiquer le passé ?			
Jamais	Une fois par mois	Au moins une fois par semaine	Tous les jours

A quelle fréquence travailliez-vous avec les enfants sur les moyens d'indiquer le futur ?			
Jamais	Une fois par mois	Au moins une fois par semaine	Tous les jours

A quelle fréquence travailliez-vous avec les enfants sur les verbes irréguliers du français les plus fréquents (être, avoir, faire, dire, aller) au présent ?

Jamais	Une fois par mois	Au moins une fois par semaine	Tous les jours
--------	-------------------	-------------------------------	----------------

A quelle fréquence travailliez-vous avec les enfants sur les verbes irréguliers du français les plus fréquents (être, avoir, faire, dire, aller) au futur ?

Jamais	Une fois par mois	Au moins une fois par semaine	Tous les jours
--------	-------------------	-------------------------------	----------------

A quelle fréquence travailliez-vous avec les enfants sur les verbes irréguliers du français les plus fréquents (être, avoir, faire, dire, aller) au passé composé ?

Jamais	Une fois par mois	Au moins une fois par semaine	Tous les jours
--------	-------------------	-------------------------------	----------------

A quelle fréquence travailliez-vous avec les enfants sur le classement des mots par catégorie (animaux, légumes, fruits, arbres, etc.) ?

Jamais	Une fois par mois	Au moins une fois par semaine	Tous les jours
--------	-------------------	-------------------------------	----------------

A quelle fréquence travailliez-vous avec les enfants sur l'apprentissage du nom des lettres de l'alphabet ?

Jamais	Une fois par mois	Au moins une fois par semaine	Tous les jours
--------	-------------------	-------------------------------	----------------

A quelle fréquence travailliez-vous avec les enfants sur l'apprentissage du son des lettres de l'alphabet ?

Jamais	Une fois par mois	Au moins une fois par semaine	Tous les jours
--------	-------------------	-------------------------------	----------------

A quelle fréquence travailliez-vous avec les enfants sur l'apprentissage de l'écriture des lettres de l'alphabet ?

Jamais	Une fois par mois	Au moins une fois par semaine	Tous les jours
--------	-------------------	-------------------------------	----------------

A quelle fréquence travailliez-vous avec les enfants sur les sons élémentaires de la langue orale (phonèmes) ?

Jamais	Une fois par mois	Au moins une fois par semaine	Tous les jours
--------	-------------------	-------------------------------	----------------

A quelle fréquence travailliez-vous avec les enfants sur les relations graphème-phonème les plus fréquentes et consistantes ?

Jamais	Une fois par mois	Au moins une fois par semaine	Tous les jours
--------	-------------------	-------------------------------	----------------

A quelle fréquence travailliez-vous avec les enfants sur les relations graphème-phonème dépendantes du contexte ?

Jamais	Une fois par mois	Au moins une fois par semaine	Tous les jours
--------	-------------------	-------------------------------	----------------

A quelle fréquence travailliez-vous avec les enfants sur les relations graphème-phonème pour les marques morphologiques (différences entre écrit et oral) ?

Jamais	Une fois par mois	Au moins une fois par semaine	Tous les jours
--------	-------------------	-------------------------------	----------------